The influence of HIV on STI epidemics:

the re-emergence of lymphogranuloma venereum

Minttu Minna-Maarit Ronn

PhD thesis submitted for examination
May 2013

Department of Infectious Disease Epidemiology
Imperial College London

Declaration

I declare that the work presented in this thesis is the result of original research carried out by the author Minttu Rönn. All sources used in this thesis have been cited appropriately.

This thesis has been completed under the supervision of Professor Helen Ward and Dr Peter White at the Department of Infectious Disease Epidemiology, Imperial College London, and Dr Gwenda Hughes at STI Section at the Centre for Infectious Disease Surveillance and Control in Colindale, Public Health England.

Signed: ----------

Minttu Rönn

Abstract

Lymphogranuloma venereum (LGV) is a re-emerging sexually transmitted infection (STI) in men who have sex with men (MSM). The emergence of LGV has occurred in parallel to an increase in STIs in MSM. It represents a public health problem and an added burden in the control of STIs. This thesis aims to identify factors that have contributed to the persistent high levels of HIV prevalence among LGV cases. I use surveillance data to explore the re-emergence of LGV through statistical analyses. I then perform a literature review to better understand the social epidemiological context in which seroadaptive behaviours in HIV-positive MSM occur, a hypothesised cause for the re-emerging STIs. I finally look at the dynamics of HIV and LGV through a deterministic mathematical model where the effect of serosorting is investigated in more detail.

Through statistical analyses I demonstrate that there is a strong behavioural component in explaining the association between the two infections. In the LGV Enhanced Surveillance HIV-positive LGV patients were more likely to report unprotected receptive anal intercourse compared to HIV-negative/unknown LGV patients. In a subsequent analysis I show that individuals with reported LGV re-infection were more likely to be HIV-positive, visit a clinic in London and have concurrent hepatitis C and gonorrhoea on their first recorded LGV episode. However, the data also suggests there is a diagnostic bias in favour of HIV-positive men who present with shorter duration of symptoms. I propose a conceptual framework of seroadaptive behaviours in HIV-positive MSM based on a literature review. The mathematical model demonstrates that infections with the same mode of transmission will be associated, but it provides further support to the hypothesis that serosorting can explain the particularly high HIV prevalence in those who acquire LGV.

This work cannot exclude the possibility of a biological interaction but the current evidence points to behavioural, and likely network-level, differences between HIV-positive and -negative MSM as the main driver in LGV re-emergence. This has implications for surveillance and control of LGV. Rare STIs, such as LGV, may benefit from a more detailed and qualitative investigation of cases using a periodic sentinel model as well as from health promotion targeted specifically to the small subpopulation affected.

Acknowledgements

I would like to begin the acknowledgments by thanking my main supervisor Professor Helen Ward for the long journey we have had with LGV (it certainly feels like that now!). Thank you for the many discussions, making me question the dogmas in the field and for allowing me to search for my own path and answers during my studies. I would also like to thank my supervisor Dr Peter White for the fun conversations and for making mathematical modelling make sense. I am grateful for the collaboration with Public Health England which allowed me to access the surveillance data. For this I would like to thank my supervisor Dr Gwenda Hughes whose help has been very valuable for me during this process.

There are also other people at the Public Health England who have helped me during my PhD studies, and whose contribution I would like to acknowledge: Professor Cathy Ison, Dr Sarah Alexander and Dr Ian Simms.

There are some truly wonderful people who have improved the quality of this work and my life during the write-up: Annick Borquez is now an LGV expert, and she deserves a special recognition for having gracefully shared office space with me for the past couple months. Her ability to bring sense and direction into moments of modelling confusion and delightful stories during the work day have kept me motivated, and made work fun. Jocelyn Elmes and Hannah Clapham have read and commented chapters with astonishing speed, detail and insight. I could not be more appreciative for the help these ladies have given me. This paragraph would not go without a mention of James Truscott who has taught me statistics, mathematical modelling and good practice in quantitative research amidst eating cookies and commenting on the state of the world. I would also like to thank Bethan Davies and Tanvi Rai for the peer support (and many chats) we have had during my PhD.

The lovely people at the Department of Infectious Disease Epidemiology have caused me to have too long coffee breaks, go for lunch-time runs in Hyde Park, for drinks after work and climb Snowdon in horrendous weather. There are many people here who have become good friends to me over the years.

I owe the greatest thanks to my family. My mother Marja-Leena Rönn and my brother Johann Rönn have seen me go back and forth Finland and abroad for almost a decade now. I thank you for letting me have my adventures and never questioning the sensibility of my endeavours. I thank my grandparents, Bertta and Aarno Korhonen, for being a safe haven in this world for as long as I can remember. Their kitchen dinner table has always been the place I most miss when I have been away. I also thank my aunt, Maarit Korhonen, for pushing me to test my boundaries and making me work harder. It has got me here.

Table of contents

List of figures

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

List of tables

Acronyms

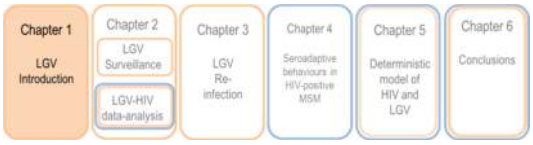
ART	Antiretroviral treatment
BASHH	British Association for Sexual Health and HIV
CI	Confidence interval
CrI	Credibility interval
ESSTI	European Surveillance for STIs
GRASP	Gonococcal Resistance to Antimicrobials Surveillance Programme
GUD	Genital ulcer disease
GUM	Genitourinary medicine
GUMCAD	Genitourinary Medicine Clinic Activity Dataset
HAART	Highly active antiretroviral treatment
HIV	Human immunodeficiency virus
HPA	Health Protection Agency
LGBT	Lesbian, gay, bisexual and transgender
LESS	London Enhanced Syphilis Surveillance
LHS	Latin Hypercube Sampling
LGV	Lymphogranuloma venereum
MSM	Men who have sex with men
NAAT	Nucleic acid amplification test
Natsal	National Survey of Sexual Attitudes and Lifestyles
NESS	National Enhanced Syphilis Surveillance
NHS	National Health Services
ONS	Office of National Statistics
ompA	Outer membrane protein A (of *Chlamydia trachomatis*)
OR	Odds ratio
PCR	Polymerase chain reaction
PEP	Post-Exposure Prophylaxis
PHE	Public Health England (formerly Health Protection Agency)
PrEP	Pre-Exposure Prophylaxis
R_0	Basic reproductive number
RITA	Recent Infection Testing Algorithm
RFLP	Restriction fragment length polymorphism
SI	Susceptible-infectious
SIS	Susceptible-infectious-susceptible
SOPHID	Survey of Prevalent HIV Infections Diagnosed
STBRL	Sexually Transmitted Bacteria Reference Laboratory
STI	Sexually transmitted infection
UAI	Unprotected Anal Intercourse
WHO	World Health Organisation

Chapter 1

Introduction

The re-emergence of lymphogranuloma venereum

1.1 Summary



This chapter provides an introduction to the epidemiology of lymphogranuloma venereum (LGV), which is a sexually transmitted infection (STI) caused by a variant of *Chlamydia trachomatis*. LGV is a re-emerging STI among men who have sex with men (MSM) in whom LGV outbreaks have been reported from early 2000s onwards in high-income countries. LGV re-emergence has been characterised by atypical clinical manifestation, the need for specialised diagnostics and new surveillance systems to monitor the occurrence of the disease, posing new challenges to sexual health services and STI research.

The emergence of LGV has occurred along with an increase in other STIs in MSM. LGV cases identified in affected countries have been remarkably similar, with the vast majority of LGV seen in HIV-positive MSM who report high levels of risk behaviour. There is a strong association between HIV and LGV which was quantified in a previous meta-analysis of four case-control studies where LGV patients were over eight-times more likely (OR 8.2, 95% CI 4.7, 14.3) to be HIV-positive in comparison to controls with non-LGV chlamydia.

This thesis aims to understand the relationship between HIV and LGV epidemics from an epidemiological and public health perspective. The analyses presented can hopefully contribute to LGV surveillance and control methods and to further characterise risk among MSM populations, and HIV-positive MSM in particular. HIV and STI prevention has traditionally focused on the HIV-negative but with an increasing pool of HIV-positive individuals and increasing levels of sexual risk behaviour reported among MSM, regardless of serostatus, involving HIV-positive individuals into prevention programmes becomes crucial.

1.2 Lymphogranuloma venereum

LGV is a sexually transmitted infection (STI) which re-emerged in high-income countries among men who have sex with men (MSM) in the early 2000s. The emergence was characterised by the majority of cases being diagnosed in HIV-positive MSM, a delayed recognition of LGV due to its rarity and a clinical presentation which was unlike that seen in text books (White & Ison, 2008). The modes of transmission, risk factors for infection and natural history of LGV are still not well understood (de Vries, van der Bij, Fennema, *et al.*, 2008; Hamill, Benn, Carder, *et al.*, 2007).

LGV was first noticed in an HIV-positive gay man in Rotterdam (Nieuwenhuis, Ossewaarde, van der Meijden, *et al.*, 2003)[1]. Following a cluster of LGV (Götz, Nieuwenhuis, Ossewaarde, *et al.*, 2004), an alert of LGV outbreak was announced by the European network for surveillance of STIs (ESSTI) (von Holsterin, Fenton & Ison, 2004). Case-reports began to emerge from Europe, North America and Australia[2]. From these early outbreaks, LGV has now established as an endemic infection among MSM in large European metropolitan areas. The most comprehensive surveillance data comes from United Kingdom (UK) and the Netherlands, with the UK having the largest documented outbreak of LGV (Ward, Alexander, Carder, *et al.*, 2009; de Vrieze, van Rooijen, van der Loeff, *et al.*, 2013). MSM have been disproportionately affected by the HIV epidemic across the world (Beyrer, Baral, van Griensven, *et al.*, 2012), and LGV is one of many STIs on the increase (Fenton & Imrie, 2005) which has implications for surveillance and control of STIs. The surveillance data for LGV is managed by Public Health England (PHE) and through collaboration with their STI Section at the Centre for Infectious Disease Surveillance and Control in Colindale[3], I have been granted access to the LGV Enhanced Surveillance data which offers an opportunity to improve our understanding of the LGV re-emergence in MSM.

[1] The first case report was later linked to controversy, when it was revealed that formal notification of the episode was delayed until the case-report was published (Vermij, 2005).

[2] (von Holsterin, Fenton & Ison, 2004; Vall Mayans, Sanz Colomo & Ossewaarde, 2005; Plettenberg, von Krosigk, Stoehr, *et al.*, 2004; Berglund, Bratt, Herrmann, *et al.*, 2005; Liassine, Caulfield, Ory, *et al.*, 2005; Ahdoot, Kotler, Suh, *et al.*, 2006; Kropp, Wong & Canadian, 2005; Morton, Fairley, Zaia, *et al.*, 2006)

[3] Previously known as the HIV and STI Department of Health Protection Agency.

In the following section I will give an introduction to LGV by describing its biology, diagnostics and treatment followed by a characterisation of the association between HIV and LGV. The introduction builds on previous work which involved a systematic review and meta-analysis of LGV re-emergence (Rönn, 2009).

1.2.1 Biology and natural history

LGV is a subtype of the bacterium species *Chlamydia trachomatis* which is part of the order Chlamydiales. Chlamydiales are distinguished by their unusual lifecycle involving intracellular replication in the host due to their inability to biosynthesise necessary macromolecules. The species in the order are obligate pathogens and they have a host range of different vertebrates including humans, birds and domestic animals. The diminished genome of *C. trachomatis* further reflects its dependence on host metabolic functions[4] (Schachter and Stephens, 2008).

C. trachomatis has a number of serovars which represent minute genetic differences yet cause three different diseases (Thomson et al., 2008): serovars A-C are the cause of trachoma, an eye infection that mainly spreads via child-to-child transmission, serovars D-K cause the "common" anogenital chlamydia whilst serovars L1, L2 and L3 (and their variants) cause LGV (Stamm, 2008b). The natural history of LGV is poorly understood. The frequency of infection following exposure is unclear but LGV is probably not as contagious as gonorrhoea (Stamm, 2008b; Mabey & Peeling, 2002). In mixed chlamydial infections with both non-LGV and LGV serovars, it has been shown that LGV has a 10-fold lower bacterial titre compared to non-LGV chlamydia indicating greater fitness in non-LGV chlamydia (Morre, Ouburg, van Agtmael, *et al.*, 2008). Based on biological plausibility the primary ulcerative skin lesions, urethritis, cervicitis, proctocolitis and chronic ulceration are likely to be most infectious for LGV (Stamm, 2008b).

The life cycle of *C. trachomatis* consists of an extracellular elementary body, which is not metabolically active and is relatively resistant to the extracellular environment, and an intracellular reticulate body responsible for

[4] LGV was thought to be caused by a lymphotropic virus still in the 1950s (Coutts, 1950). Due to the intracellular replication and relative infrequency of co-infections with multiple strains, recombination was previously considered rare in chlamydia but instead has been recently shown to be a relatively common event that has occurred also between the ocular and anogenital *C. trachomatis* strains (Harris et al. 2012). LGV-D chlamydia recombinant has also been identified in the current emergence(Somboonna, Wan, Ojcius, *et al.*, 2011).

the replication of the organism. The elementary body attaches to its target cell and is engulfed by the host cell by endocytosis and remains in an endosome whilst in the host cell. There, the pathogen undergoes a morphological change into a reticulate body which is capable of intracellular growth and replication. The intracellular cycle takes place in a chlamydia-modified vacuole which the pathogen creates by inducing phagolysosomal fusion in the host cell. The reticulate body divides for 20-24 hours post-entry after which the number of elementary bodies in the vacuole begins to increase. Forty-eight to 72 hours post-entry the intracellular vacuole and then the cell ruptures and elementary bodies are released. Non-LGV strains of *C. trachomatis* have a limited host-cell range in vivo, and are mainly capable of infecting squamocolumnar epithelial cells, while LGV is less tissue specific and is also capable of infecting lymphatic tissue (Schachter & Stephens, 2008).

C. trachomatis, presumably including LGV, cannot penetrate intact skin and requires small abrasions and lesions to gain entry into the host. LGV is able to infect macrophages and spread to the lymphatic tissue at the site of infection leading to a more systemic disease in comparison to non-LGV chlamydia. There is some evidence to support partial immunity to chlamydial infection resulting from repeated exposure, but no LGV-specific data exist (Brunham, 2013; Stamm, 2008a).

1.2.2 Clinical manifestation

In its initial phase LGV can present as a genital ulcer disease (GUD) (Kaliaperumal & Karthikeyan, 2008) and the manifestation can be similar to genital herpes and chancroid. Durand, Nicholas and Favre established LGV as a clinical entity in 1913 (Stamm, 2008b) and it has been known with many names (climatic bubo, Buba, Nicolas-Favre's disease, chronic elephantiasis, 4^{th} and 6^{th} venereal disease, lymphogranuloma inguinale) (reviewed by Coutts, 1950; Annamunthodo, 1961). There are earlier descriptions of lymphadenitis (inflammation of lymph nodes) by the Greek, Roman and Arab physicians that were indicative of LGV (Richardson & Goldmeier, 2007), and “bubo” was reported to be common in Medieval Italy.

As the name lymphogranuloma venereum implies, LGV infection involves inflammation of the lymph nodes. Given the lack of research into LGV prior to the current re-emergence, relatively little is known of the natural history of the infection. Textbooks have divided LGV disease progression into three stages with the primary

stage involving the site of inoculation where a small ulcer, papule, lesion or nonspecific urethritis appears following an incubation period of 3-30 days. The secondary stage can begin from 10-30 days to months later. During the secondary stage lymph nodes at the site of infection enlarge as a result of an inflammatory process resulting in chronic oedema, ulceration and sclerosing fibrosis. In men LGV has been reported to classically present as inguinal lymphadenopathy (enlarged lymph nodes at groin area), which assumes the urethra as the site of infection (Stamm, 2008b; Coutts, 1950). In women the primary site of infection is more often in the vagina, rectum, cervix or posterior urethra. These drain to the deep iliac or perirectal lymph nodes, and present less often with clear clinical signs of early LGV infection (Roest, van der Meijden, European Branch of the International Union against Sexually Transmitted, *et al.*, 2001) and only 20-30% are said to experience inguinal syndrome. LGV has also been documented to present as acute systemic infection with fever without clear lymph node or tissue reaction at the infection site (Stamm, 2008b). It has also been suggested that latent LGV and spontaneous recovery are possible following the secondary stage (Stamm, 2008b; Clinical Effectiveness Group of the British Association for Sexual & HIV, n.d.; Kaliaperumal & Karthikeyan, 2008), but there is little evidence of any of the possible stages in the current LGV epidemic. In the absence of treatment a tertiary stage with a chronic inflammatory response follows. This is characterized by genital ulcers, fistulas, rectal strictures and genital elephantiasis. The scarring and formation of fibrotic tissue often require surgical repair (Stamm, 2008b; Coutts, 1950). In a study of MSM presenting to surgical practice due to proctitis-related (inflammation of rectum) sequel 4 out of 26 had LGV (Davis & Goldstone, 2009).

In the current emergence of LGV, the rectum has been the most common site of infection with acute ulcerative proctitis or proctocolitis (inflammation of rectum and colon) often seen as the primary signs in LGV patients (White, 2009). Approximately 96% of the episodes seen in the UK have been rectal LGV (Ward, Martin, Macdonald, *et al.*, 2007) whilst inguinal lymphadenopathy alone has been observed only in a few cases (Sethi, Allason-Jones, Richens, *et al.*, 2008). The symptoms most frequently reported are rectal discharge, pain, bleeding (White, 2009) as well as systemic symptoms (Macdonald, Ison, Martin, *et al.*, 2005) and high white blood cell counts (Van der Bij, Spaargaren, Morre, *et al.*, 2006).

LGV re-emergence has been described as having an atypical clinical manifestation when compared with accounts in the modern medicine textbooks. However, rectal LGV was not uncommon in the older medical literature (Annamunthodo, 1961; Miles, 1957; Coutts, 1950): rectal LGV was seen as a sequel of primary LGV (as the infection progressed through the lymphatic tissue over a long duration of infection) or the first noticeable symptom if the primary site of infection was not connected to visible inguinal lymph nodes. Annamunthodo (1961) describes three varieties of LGV: inguinal LGV (with inguinal lymphadenopathy), genital LGV (with urethral involvement) and anorectal LGV (described by the lack of acute phase followed by proctitis which has spread from a nearby site, such as the cervix, or by direct inoculation into the rectum).

While a majority of the LGV cases found in a large case-finding exercise in the UK have been symptomatic (Ward, Alexander, Carder, *et al.*, 2009), in the Netherlands almost half of the cases have been asymptomatic (Van der Bij, Spaargaren, Morre, *et al.*, 2006; Spaargaren, Fennema, Morre, *et al.*, 2005; de Vrieze, van Rooijen, van der Loeff, *et al.*, 2013). It has been suggested that time between becoming infected and seeking treatment would explain the asymptomatic and symptomatic cases (White, 2009).

1.2.3 Diagnosis and treatment

LGV diagnosis was first based on clinical symptoms and after 1925 has relied on antibody recognition when Frei's test was introduced (skin test similar to tuberculin test) (Kapoor, 2008; Coutts, 1950). Later more refined techniques have been developed with enzyme immunoassay and complement fixation being the most typical. Antibody-based tests are prone to cross-reactivity with other *Chlamydia trachomatis* serovars and potentially also with other species of the Chlamydia genus; they also require a strong immune response to create a positive result, thus contradictory evidence exists of their usefulness in LGV diagnostics (van der Snoek, Ossewaarde, van der Meijden, *et al.*, 2007b; Smelov, Morre, de Vries, *et al.*, 2008; van der Snoek, Ossewaarde, van der Meijden, *et al.*, 2007a).

The routine chlamydia test does not differentiate between LGV and non-LGV chlamydia and in the context of the current re-emergence the LGV case definition requires confirming the presence of LGV serovar (L1-L3) in the sample. Most typically this occurs by identifying the *ompA* (outer membrane protein A) gene variant which is well characterized for *C. trachomatis* and differs for each serovar (Thomson, 2008). Polymerase chain

reaction (PCR) using restriction fragment length polymorphism (RFLP) is the most common method though genotyping is also used. Despite the developments, LGV diagnosis requires an additional step after a sample has been identified as chlamydia positive (Morre, Ouburg, van Agtmael, *et al.*, 2008). After the re-emergence of LGV, specific PCR tests have been developed to facilitate faster detection that also allows identifying mixed chlamydial infections (Morre, Spaargaren, Fennema, *et al.*, 2005; Halse, Musser & Limberger, 2006; Jalal, Stephen, Alexander, *et al.*, 2007; Chen, Chi, Alexander, *et al.*, 2007). However, the US Food and Drug Administration is yet to approve STI detection from rectal samples using commercially available tests, despite the demonstrated high sensitivity and specificity and a need to test non-genital sites. This limits the numbers of laboratories offering chlamydia testing from rectal and pharyngeal sites (Alexander, 2009; Cosentino, Campbell, Jett, *et al.*, 2012). LGV confirmation remains restricted to specialised laboratories (Morre, Ouburg, van Agtmael, *et al.*, 2008).

Non-LGV chlamydia is normally treated with a single dose (1g) of azithromycin (or alternatively 7 days of doxycyclin) but with LGV a prolonged treatment with doxycycline for 21 days is currently recommended, as treatment failures have been reported with the standard chlamydia treatment. Also a test of cure is recommended (de Vries, Smelov, Middelburg, *et al.*, 2009; McLean, Stoner & Workowski, 2007).

1.2.4 LGV epidemiology

1.2.4.1 Geographical distribution of heterosexual LGV

LGV is reported to be endemic in East and West Africa, Southeast Asia and India, and Caribbean and South America (Stamm, 2008b). To examine the epidemiology of heterosexual LGV there are more studies from areas which were traditionally perceived as the endemic regions for LGV: in South Africa the proportion of LGV from genital ulcer diseases ranges between 3-8.4% in women and 7-17.4% in men based on a review (Johnson, Coetzee & Dorrington, 2005) whilst O'Farrell et al. (2008) found that 13.2% of GUDs in men were LGV (with 86.4% HIV co-infection). Wide variation of LGV is reported from India: a decline (3.4% to 0.2% of GUDs) in LGV cases was reported in North India between 1990-2004 while 9.3% of STI cases (51/551) in an Air Force hospital were LGV (Jaiswal, Banerjee, Matety, *et al.*, 2002) and in Kolkata 0.47% of diagnosed STIs were LGV (Ghosh, 2002). A Kuwaitian study found LGV in 1.4% of STI cases (Al-Mutairi, Joshi, Nour-Eldin, *et al.*, 2007). In

these studies LGV has been diagnosed by clinical symptoms or by enzyme immunoassay and given few countries require official notification prevalence of LGV in these areas remains uncertain.

During the current re-emergence few studies have focused on LGV in heterosexuals in the high-income countries: Switzerland (Goldenberger, Dutly & Gebhardt, 2006), France (Herida, Kreplack, Cardon, *et al.*, 2006) and Spain (with one LGV case found in a homosexual man) (Pineiro, Montes, Gil-Setas, *et al.*, 2009) found no LGV in the samples, and in the absence of evidence to the contrary LGV is believed to be confined to MSM.

LGV became rare in the industrialised world after the introduction of antibiotics with a few dozen cases reported in Europe annually since 1950s (Stamm, 2008b). During the last decades there were few published outbreaks of LGV with 27 cases in Paris in 1989, seven cases in Seattle in 1995 (Richardson & Goldmeier, 2007) and 15 cases in Bahamas in 2002 among crack cocaine users (Bauwens, Orlander, Gomez, *et al.*, 2002).

1.2.4.2 LGV epidemiology in MSM

In the current re-emergence of LGV, the majority of cases have been HIV-positive MSM with rectal infection. Other STI co-infections have been common, most notably 19% of cases during early epidemic in the UK also had sexually acquired hepatitis C infection (Ward et al., 2007). LGV cases have also been older (with mean age above 35) compared to MSM presenting to STI clinic for other reasons (van de Laar, 2006).

In a systematic review performed earlier (Rönn, 2009) few studies were found to estimate the burden of LGV among MSM. Based on two studies conducted in London and Brighton (Ward, Alexander, Carder, *et al.*, 2009; Annan, Sullivan, Nori, *et al.*, 2009) a joint estimate of LGV positivity in rectal samples of MSM attending GUM clinics was 0.98% (95% CI 0.8, 1.2). Prevalence of rectal chlamydia was estimated in these two studies and a third LGV study from Swedish clinic attendees (Cullberg, Bratt, Petrersson, *et al.*, 2009) resulting in a pooled estimate of rectal chlamydia positivity in the overall sample 8.5 % (95% CI 8.0, 9.0) and 4.3% (95% CI 3.9, 4.7) for urethral chlamydia. Urethral LGV was looked at in the case-finding exercise by (Ward, Alexander, Carder, *et al.*, 2009) giving an estimate of 0.04% for urethral LGV positivity. The study by Annan was the only to estimate overall LGV among MSM at 1.2% (95% CI 0.8, 1.6).

The LGV epidemic curve in the United Kingdom is shown in Figure 1 where laboratory confirmed episodes of LGV are presented since the beginning of LGV surveillance in 2004. Although the number of cases has remained low in general, and the infection seems to remain confined in small high-risk groups, the overall number of diagnosed LGV cases has been on the increase since the beginning of surveillance, and it has been able to establish itself in a subpopulation of MSM.

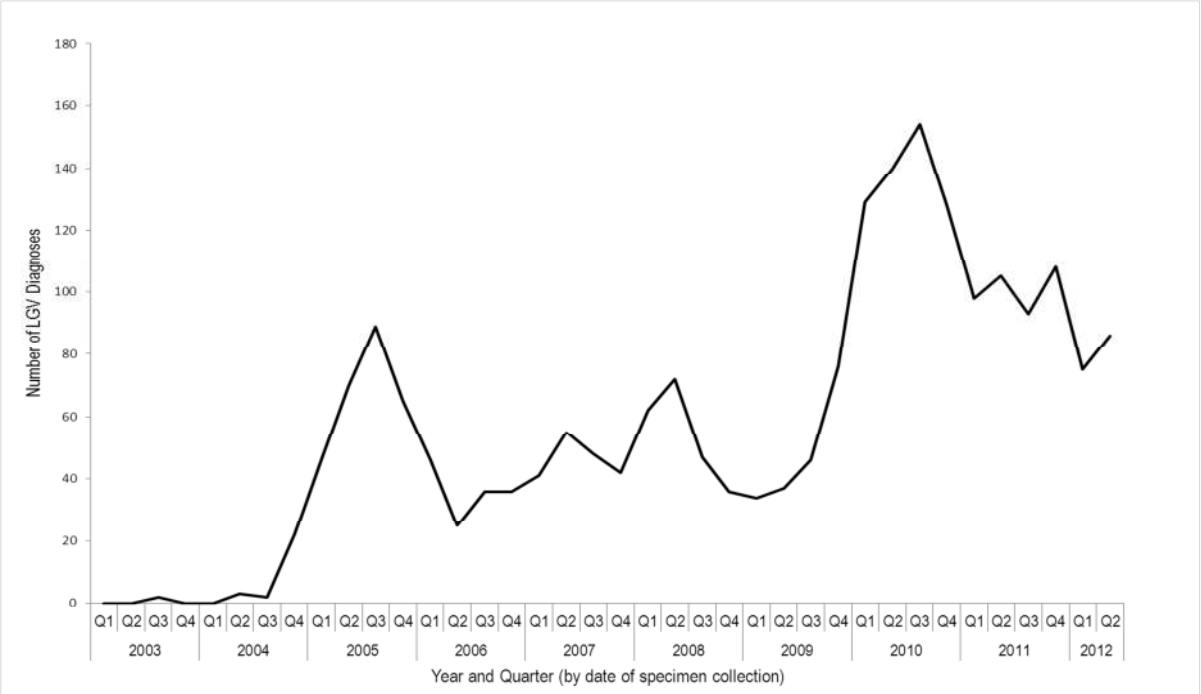


Figure 1. Number of laboratory confirmed LGV cases in the United Kingdom. Figure obtained from the Public Health England.

It is not clear why urethral infections have been uncommon, as it would imply rectum-to-rectum transmission to be common. It has been suggested that different modes of transmission, such as sexual practices of fisting and use of sex toys and other fomites, and practice of enema use are contributing to the spread of LGV (de Vries et al., 2008, Hamill et al., 2007). But these are also likely to act as confounders or intermediate factors in most instances with unprotected anal sex being a likely source of infection in most cases (Ward, Macdonald, Ronn, *et al.*, 2011). Traumatic practices to the mucosa have been associated with the acquisition of other rare STIs such as sexually acquired hepatitis C virus (Danta, Brown, Bhagani, *et al.*, 2007). Non-LGV *C. trachomatis* serovars exhibit tropism for different epithelial cell types and there is larger variation in chlamydia serovars

within heterosexual populations compared to MSM. Furthermore sexual networks of MSM demonstrate more clustering of strains with a larger overlap between *C. trachomatis* serovars across countries in comparison to heterosexuals (Jeffrey, Suchland, Quinn, *et al.*, 2010; Christerson, Bom, Bruisten, *et al.*, 2012).

1.2.5 Is this an outbreak or a re-discovered endemic disease?

It has been suggested that the re-emergence of LGV in MSM could be an artefact of improved diagnosis rather than a true outbreak (Schachter & Moncada, 2005). It is true that the ability to detect LGV has improved with the implementation of NAATs and screening of rectal STIs has increased in STI clinics serving large MSM populations (Alexander, 2009). The lack of standardised surveillance systems prior to the recognition of the increase in LGV cases makes it difficult to estimate how much the re-emergence is due to increase in incidence compared to infection being detected more often (Pathela, Blank & Schillinger, 2007). Retrospective studies have identified LGV in stored specimens (Gebhardt & Goldenberger, 2005; Waalboer, van der Snoek, van der Meijden, *et al.*, 2006; Herida, Sednaoui, Couturier, *et al.*, 2005; Halioua, Bohbot, Monfort, *et al.*, 2006), most notably from specimens in San Francisco in the 1980s (Spaargaren, Schachter, Moncada, *et al.*, 2005) and LGV may have been overlooked until changes in sexual behaviour and the emergence of seroadaptive behaviours in MSM facilitated the establishment of LGV among other STIs [5].

However, recent evidence points to the direction that the LGV strain circulating in the MSM is a result of a clonal expansion (Harris, Clarke, Seth-Smith, *et al.*, 2012): in a whole-genome sequencing of 36 *C. trachomatis* genomes L2b serotype had a very low level of variation with a maximum pairwise evolutionary distance of 19 single nucleotide polymorphisms between the 12 L2b strains analysed from the global collections. The authors infer from their results that this could be due to fast transmission and compare this to the Swedish chlamydia variant which spread rapidly in the population due to deletion in coding sequence used by detection in NAATs (Herrmann, Eden, Hadad, *et al.*, 2012). Harris *et al.* (2012) suggest that these chlamydia strains demonstrate

[5] This makes an interesting counterpart to an old World Health Organisation (WHO) review on LGV (Coutts, 1950), which comments on LGV epidemiology much the same way as we do today: LGV was then considered a disease of tropical regions but accurate information on prevalence was not available. Furthermore due to recently developed Frei's test, clinicians were suddenly identifying LGV in every country the test was in use. (Given the lack of sensitivity or specificity of Frei's test, this makes a historic anecdote of epidemiological investigation rather than epidemiological evidence in itself.)

that given a selective advantage and/or lack of competition, a strain of *C. trachomatis* can emerge relatively quickly.

This hypothesis is further supported by the geographical clustering of LGV cases (Jebbari, Alexander, Ward, *et al.*, 2007) and the small number of cases seen in high-risk MSM would indicate LGV has found a population with behavioural patterns and international sexual networks that facilitate the establishment of the disease (van de Laar, 2006) similar to that observed with sexually acquired hepatitis C in MSM (van de Laar, Pybus, Bruisten, *et al.*, 2009).

It is noteworthy that there are areas where the LGV re-emergence can be seen as a true outbreak: Sweden began LGV surveillance in 2004 after two imported cases were identified but no cases were found in MSM with chlamydia (the study looked at 81% of chlamydia episodes reported in MSM in a 13 month period) (Klint, Lofdahl, Ek, *et al.*, 2006). Three years later in 2007, 15 LGV cases were diagnosed, with some reporting Sweden as the probable country of acquisition (Velicko, Cullberg, Bratt, *et al.*, 2009). Similarly Canada (Tinmouth, Gilmour, Kovacs, *et al.*, 2008) and Australia (Lister, Tabrizi, Fairley, *et al.*, 2004) have performed case-finding exercises where the countries did not find LGV in their sample but both countries have subsequently seen LGV. Moreover Lima, Peru (Clark, Espinosa, Leon, *et al.*, 2008) and Alabama, USA (Geisler, Morrison & Bachmann, 2008) have reported not finding any LGV in their study sample further suggesting LGV distribution in MSM to be highly skewed (it is worth mentioning that many of these studies have had a small sample size in relation to the paucity of LGV). These different pieces of evidence would suggest the current epidemic is the result of a new outbreak rather than merely increased testing for LGV.

1.2.6 LGV as a public health problem

The emergence of LGV has involved characteristics which have made the public health response more difficult as discussed in Timen *et al.* (2008) due to lack of awareness, problems in diagnosis, limited surveillance and difficulties in partner notification. Given the rarity of LGV in industrialised world, clinicians did not consider it in differential diagnosis in the beginning of the re-emergence, and the clinical presentation has been different to the standard description of LGV, often mimicking inflammatory bowel disease and other chronic conditions. In addition, LGV is circulating in sexual networks which are characterised by anonymous partners, with

international links making contact-tracing difficult. LGV is not notifiable by law in most European countries and affected countries have been varied in their response with some establishing surveillance systems specific to LGV. Diagnosis of rectal LGV involves proctoscopy and requires nucleic acid amplification tests (NAAT) that were not licensed for rectal specimens. Furthermore, based on a GUM clinic survey, there has been an inconsistent clinical practice regarding screening MSM for rectal chlamydia and cases are likely to be missed until/unless they develop severe symptoms (McMillan, Kell, Ward, *et al.*, 2008). The establishment of LGV may have been facilitated by the inconsistent clinical practice in screening and treating for LGV in the UK and elsewhere with case definitions and surveillance methods varying between countries (van de Laar, 2006; Timen, Hulscher, Vos, *et al.*, 2008). LGV represents an aggressive infection which emerged and has established itself despite the control measures put in place. LGV is representative of other STIs on the rise in MSM, a phenomenon which has become one of the main problems for STI control in high-income countries, and it is an added burden to the health care system.

1.3 HIV and STIs in men who have sex with men

There are biological and behavioural factors that make MSM particularly vulnerable to HIV acquisition and transmission, most notably role versatility in (in contrast to the role segregation in heterosexual sex), and receptive anal sex causing a higher transmission probability in comparison to vaginal intercourse (Baggaley, White & Boily, 2010; Goodreau & Golden, 2007). The highest sexual activity group has disproportionate impact on the rest of the population in MSM populations (compared to heterosexual population with similar behaviour) as demonstrated in a mathematical modelling work by Goodreau & Golden (2007). Also the sexual mixing patterns among MSM are different compared to heterosexuals and MSM report more dissortative mixing by age, more partnership concurrency, and larger number of partners than heterosexuals (in a study conducted in the US; Glick, Morris, Foxman, et al., 2012).

HIV prevalence in MSM surpasses the general population HIV prevalence in every setting it has been measured (Beyrer, Baral, van Griensven, *et al.*, 2012). Conceptualisation of homosexual behaviour became a research interest along with the rise of the HIV epidemic, and national surveys began measuring prevalence of homosexual activities (reviewed by Michaels & Lhomond, 2006). This has primarily focused on sexual acts and

partners, with some surveys later on including sexual identity in response to lesbian, gay, bi and transgender (LGBT) mobilisation. The difference between defining and separating behaviours and identity becomes important if we are trying to understand the epidemiology of infection (behaviour of primary interest) or if we are planning health promotion and policies (how to reach the target population)[6].

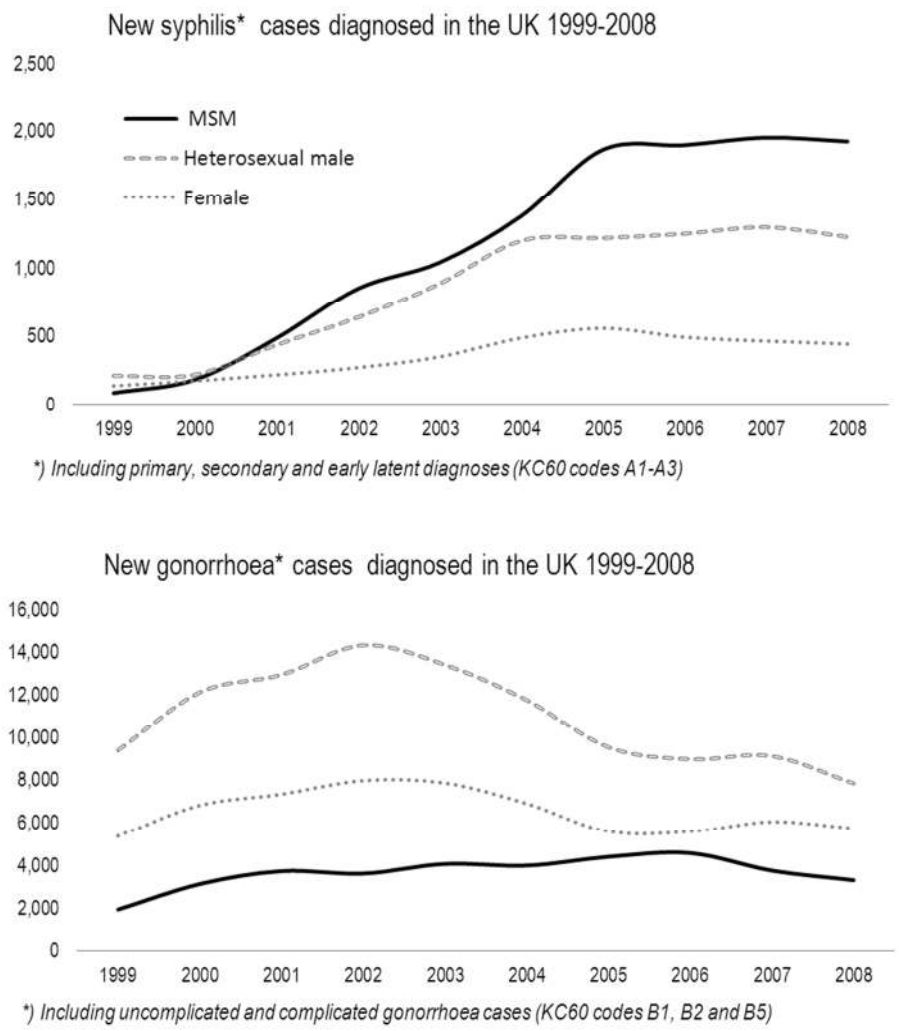


Figure 2. Illustration of a trend by smoothing over yearly number of diagnoses for syphilis and gonorrhoea, stratified by sex and sexuality. These are based on GUM clinic KC60 returns (England, Wales and Northern Ireland) and ISD(D)5 and STISS returns from Scotland. Data was obtained from Annual STI Data Tables (Public Health England, 2010d).

[6] In the UK, National Survey of Sexual Attitudes and Lifestyles (Natsal) measures population level sexual behaviours and does not ask about sexual identity (Mercer, Fenton, Copas, *et al.*, 2004), whilst Gay Men's Sex Survey is a gay community-based survey which has occurred annually since 1997 and it includes men who either have had sex with men, intent to have sex with men, or who self-identify as gay or bisexual (or other similar identity) (Sigma Research, 2010).

Over the past decade or so there has been an increase both in the number and types of STIs seen in MSM with a rise in rates of gonorrhoea, syphilis, hepatitis C and outbreaks of enteric pathogens such as shigella. Number of syphilis and gonorrhoea diagnoses in the UK is illustrated in Figure 2. Syphilis diagnoses substantially increased in the United Kingdom between 2000 and 2005 parallel to the increase seen in Europe, and North America (Fenton, 2004). Prior to the outbreak syphilis cases had remained low between 1988 and 1996, and syphilis last peaked in the UK in the late 1970s. Syphilis re-emerged in Bristol in 1997 among heterosexuals and subsequently spread to other cities in the country and localised outbreaks have been identified in Manchester, Brighton and London (Simms, Fenton, Ashton, *et al.*, 2005). The syphilis epidemic has since become more concentrated among MSM than heterosexuals and 73% of episodes reported to National Enhanced Syphilis Surveillance (1998-2008) have been in MSM of which 35% of MSM have been HIV-positive (Jebbari, Simms, Conti, *et al.*, 2011).

While there are still more gonorrhoea cases occurring in heterosexuals the incidence is higher in MSM as shown in a study conducted in London where gonorrhoea incidence was strongly correlated with black ethnicity among heterosexuals, and with being MSM (Risley et al., 2007). Furthermore gonorrhoea strains with reduced antimicrobial sensitivity have increased in MSM (Farhi, Gerhardt, Falissard, *et al.*, 2007; de Vries, van der Helm, Schim van der Loeff, *et al.*, 2009).

Sexually acquired hepatitis C is still rare but there has been an increase in MSM in high income countries since the early 2000 as reviewed by Bradshaw et al. (2013). More atypical sexually transmitted infections have also been identified in MSM: shigella (Marcus, Zucs, Bremer, *et al.*, 2004) and hepatitis A outbreaks have been identified which have predominantly affected MSM (Sfetcu, Irvine, Ngui, *et al.*, 2011).

Increased testing is likely to explain some of the trend as was noted in a cohort of HIV infected patients where implementation of routine STI screening resulted in doubling of STI diagnoses in MSM (Brook, McSorley & Shaw, 2013). The authors speculate the systematic manner of proactive STI screening implemented in the NHS trust as part of HIV patient care is largely responsible for the increase seen in number of diagnoses. Furthermore the clinical guidelines on STI testing in MSM have changed over time with an increasing

recognition and testing of STIs in extra-genital sites such as pharynx and rectum, which has created an increase in observed diagnoses in these sites (Annan, Sullivan, Nori, *et al.*, 2009; Alexander, 2009).

Reasons for the increasing trend in STIs has been explored before (Fenton & Imrie, 2005; Dougan, Evans & Elford, 2007) with increase in HIV-positive pool of MSM, increased survival due to HAART, increase in high-risk behaviour, increase in sexual market places, particularly online, have been attributed as potential reasons for the increasing rates in STIs: At the beginning of the HIV epidemic increased mortality due to HIV was attributed to decrease in STIs in MSM, such as declining syphilis diagnoses (Chesson, Dee & Aral, 2003), and this may have also changes in sexual network structures following reduction in men with high-risk sexual practices (through changes in behaviour and HIV-related mortality). This has been followed by an overall increase in sexual risk behaviours after introduction of antiretroviral therapy and replenishment of the pool of individuals engaging in high-risk behaviours (Boily, Godin, Hogben, *et al.*, 2005; Hart & Elford, 2010). In London there has been an overall increase in reported risk practices both in HIV-negative and HIV-positive MSM (Lattimore, Thornton, Delpech, *et al.*, 2011a)

Another behavioural change which has coincided with the increase in HIV-positive men is serosorting and other seroadaptive behaviours. These can be seen as harm reduction strategies, but even when HIV transmission is averted in HIV serconcordant partnerships, the potential for the transmission of other STIs remains (Fenton & Imrie, 2005; Marcus, Schmidt & Hamouda, 2011).

1.4 Association between LGV and HIV

Re-emerging infections in MSM have high co-infection with 42% HIV prevalence in syphilis in Western Europe and 32% for gonorrhoea in England and Wales (Dougan, Evans & Elford, 2007). Sexually acquired hepatitis C is mainly seen in HIV-positive MSM, with only 6 hepatitis C cases found in HIV-negative or unknown cases compared to 389 HIV-positive cases in a prospective study in London and Brighton (Giraudon, Ruf, Maguire, *et al.*, 2008).

In a systematic review and meta-analysis I previously found that countries reported a high HIV prevalence in their LGV cases, based on surveillance reports and cross-sectional studies of MSM, with 60-100% of cases occurring in HIV-positive individuals. However, there was also a significant amount of data missing on HIV status with 19.8% of all cases reported with unknown HIV status (Rönn, 2009). In four case-control studies that compared MSM with rectal LGV to MSM with rectal non-LGV chlamydia, the summary estimate for being HIV-positive was over eight-fold (OR 8.2, 95% CI 4.7, 14.3) for LGV cases compared to non-LGV cases with the forest plot shown in Figure 3 (Rönn & Ward 2011). Even though high-risk behaviour is a confounder in this unadjusted estimate, the strong association is interesting given that the control group is also composed of MSM with some risk behaviour given they have non-LGV chlamydia infection.

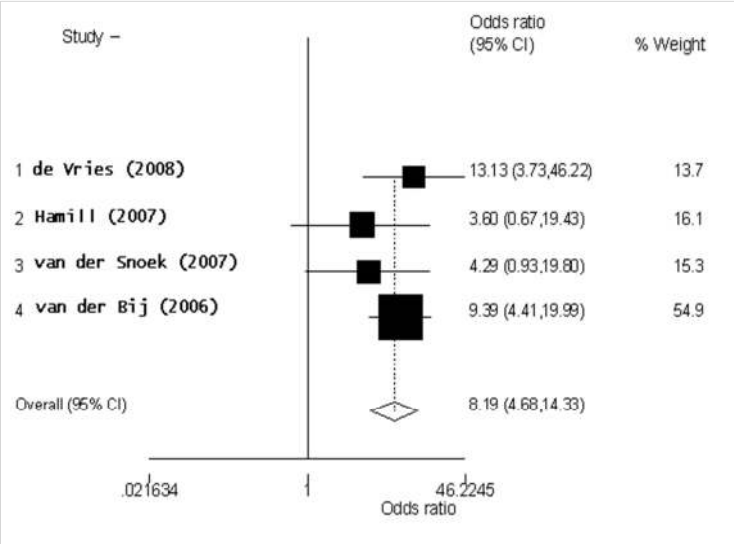


Figure 3. Forest plot of the association between LGV and HIV when LGV cases were compared to cases with non-LGV chlamydia. The figure has been published in (Rönn & Ward, 2011).

Apart from individual level risk-factors, other possible reasons for the association can be prevalence of HIV and LGV in the partner pool, serosorting and possibility for biological synergy between HIV and LGV. In an *in vitro* study of *C. trachomatis* L2 serovar and HIV-1 co-infection in epithelial cell-lines, the research group found that HIV-1 co-infection did not impact LGV replication, and although there may be indirect ways in which HIV and LGV interact , the researcher found no direct evidence of in-host interaction between the two infections (Broadbent, Horner, Wills, *et al.*, 2011).

1.5 STI and HIV transmission

There are a number of ways in which HIV and STIs can interact with each other: Underlying HIV-infection can increase the susceptibility for STI acquisition and co-infected individual can be more infectious for both the HIV and the STI. STI infection makes the person more susceptible for HIV infection through mucosal abnormalities, on-going inflammation process at the site of infection and particularly through ulcers which facilitate HIV entry (Cohen & Pilcher, 2005). HIV STI co-infection can lead to more severe progression of HIV and the severity and recovery of an STI can also be altered (progression to liver disease is more likely in hepatitis C and HIV co-infected individuals, and hepatitis C treatment success rate is lower in HIV-infected than in uninfected (Sulkowski, Mast, Seeff, *et al.*, 2000) .

In a meta-analysis of the effect of STIs on HIV susceptibility, both genital ulcer disease (GUD) and chlamydia were found to increase susceptibility to HIV-infection in heterosexuals based on evidence from observational studies (Rottingen, Cameron & Garnett, 2001). The epidemiological synergy between HIV and STIs has been extensively looked at especially in areas where the prevalence of both infections is high: among a cohort of sex workers in Kenya, being HIV-positive was associated with incident STIs with GUD having the highest hazard ratio associated with HIV-positivity (HR 2.8, 95% CI 2.0, 3.9) (McClelland, Lavreys, Katingima, *et al.*, 2005).

Based on the literature with an emphasis on papers by Rottingen, Cameron & Garnett (2001), Fenton & Imrie, (2005), Ward, Martin, Macdonald, *et al.*, (2007) and Dougan, Evans & Elford (2007) the following possible explanations for the observed association were drawn for the interaction between HIV and LGV:

- Biological difference in susceptibility
    - Increased transmission of HIV in presence of LGV
    - Increased susceptibility for LGV in HIV-positive
    - More symptomatic LGV in HIV-positive
- Diagnostic bias
    - Differences in testing and case finding
    - HIV-positive have increased contact with the healthcare system

- Difference in network structures
    - Demographic changes with larger populations of MSM and HIV-positive MSM
    - "Sexual market places"
    - Serosorting and seroadaptive behaviours
    - Increased risk behavior in certain HIV-positive sub-communities

Even though the association between STIs and HIV has been demonstrated in a number of observational settings, the practical importance of STIs to HIV transmission and the direct benefit of interventions aimed at reducing STIs in order to reduce HIV transmission at the population level have been mostly disappointing or different trials have produced contradictory results as reviewed by Ward and Rönn (2010) and Gray and Wawer (2008).

Exploring these associations is challenging: using surveillance data we cannot directly observe biological properties in relation to HIV-positivity (of the pathogens or the host in question), nor can we directly examine network structure given that information on partners and locations is collected per episode. Observational data, can be used to draw inferences on clinical and behavioural attributes on an individual level, and create hypotheses that can be tested further.

1.6 Thesis objectives and chapter plan

The re-emergence of LGV is a public health problem, and does not appear to be simply an artefact of increased testing and better diagnostic tools. LGV has established itself in the United Kingdom, and it is maintained at low levels despite additional testing and treatment set up for LGV. It is part of a wider problem of how to best control infections amidst the increasing rate of STIs in MSM.

The re-emergence of LGV offers an interesting case study of how health care systems respond to and are able to control outbreaks of STIs to which little surveillance or diagnostic capacity exist prior to the emergence. Furthermore as the burden of HIV and STIs in the Western World is concentrated in MSM, LGV also provides an interesting opportunity to look at the factors that may be contributing to the problem. There are few STIs that are so tightly confined to a specific subset of the population as LGV is to MSM with very few reports of cases among heterosexuals. For example, syphilis outbreaks in the United Kingdom began in heterosexual

population, and although syphilis has later become more common among MSM, there are still frequent cases diagnosed in heterosexuals (Simms, Fenton, Ashton, *et al.*, 2005). Another feature of LGV emergence is its association with HIV, more precisely with MSM known to be infected with HIV. Given other more common STIs, such as gonorrhoea and syphilis, have a lower level of HIV-STI co-infection, this aspect of LGV re-emergence may provide a clue to why it has been successful in reappearing in the Western world.

This thesis aims to answer some of these questions through the chapters and objectives outlined next.

In **this chapter** I have reviewed the LGV re-emergence in the light of the pathogen biology and natural history, LGV epidemiology and its association with HIV, LGV in the context of other STIs in MSM and the potential interactions of STIs and HIV.

The overall aim of this thesis is to understand the relationship between HIV and LGV and to identify the mechanisms through which these epidemics intersect. I will use statistical modelling to find associations for the measured variables in the surveillance data and mathematical modelling to explore the underlying transmission dynamics. I will also use a social epidemiological framework to examine the context in which these epidemics meet. This approach will allow combining different epidemiological methods to better understand the public health problem. The results can be used to identify appropriate responses regarding surveillance and control of LGV and to recognise further research avenues.

In **the second chapter** I will describe the profile of LGV cases in the UK, the largest documented outbreak, using laboratory and enhanced surveillance systems which have been the primary means of monitoring LGV occurrence. I will explore the profile of LGV episodes in the enhanced surveillance dataset to identify differences between HIV-positive and HIV-negative subpopulations who acquire LGV through the construction of a statistical model including clinical and behavioural factors. I will also describe of strengths and limitations of the surveillance data, assessing the utility of the systems for answering research questions.

The third chapter will focus on the traditional epidemiological concept of core groups and their importance for infection persistence in a population. LGV is hypothesised to be circulating in the “core of the core” of MSM. To explore whether core groups can be identified amongst those who have acquired LGV, a data analysis was

performed focusing on the repeat infections in the surveillance data. In this chapter, I aim to identify characteristics that might explain re-infection among the men in the dataset by comparing baseline characteristics of men with repeat LGV episodes to men with a single reported episode.

In **the fourth chapter** I will look at the population level behavioural patterns that might have facilitated LGV re-emergence by performing a literature review on seroadaptive behaviours employed by HIV-positive MSM, and proposing a theoretical framework around the theme. This provides a social context of how individual-level behaviours, embedded in social and sexual network structures, have changed in response to the HIV epidemic. The overall aim of this chapter is to better understand the social context in which seroadaptive behaviours occur. Literature review is used to develop a social epidemiological framework of factors contributing to seroadaptive behaviours and beliefs.

After the data-analyses and the conceptual framework, a population level approach will be adopted through mathematical modelling in **the fifth chapter**. The chapter investigates the conceptual processes underlying HIV and LGV transmission and a deterministic compartmental model of LGV and HIV is developed for this. I will examine to what extent behaviour is able to explain the association between the two infections. Sexual mixing patterns in the population are simulated through mixing by one's activity level and by perceived HIV status and different levels of underlying HIV prevalence in one's partner pool are investigated against LGV emergence.

Mathematical models are a useful way of presenting the system in a mechanistic way. Whilst statistical models tend to give a description of a phenomenon as observed in the data, and it quantifies the association between an exposure and outcome, it does not tell us how the underlying system works. Here mathematical modelling can formalise (and explicitly state) the causal pathways that are thought to affect the disease transmission (Garnett, Cousens, Hallett, *et al.*, 2011), which will bring an additional perspective to the framework developed in chapter four. Overall the different approaches and methods used will give a better evidence base for conclusions.

Finally, in **the sixth chapter** I will draw together conclusions, review them against other research in the field, describe the strengths and limitations of this study along with public health implications of the results, and point to further research avenues of interest.

By using different methods and theoretical perspectives, this thesis will try to disentangle the levels of influence on the intersecting epidemics of HIV and LGV as schematically illustrated in Figure 4. At the centre of the figure we have factors contributing to the transmission of HIV and LGV and on the left we have the thesis chapters pointing to areas to be covered. Of the hypotheses of reasons for the association between HIV and LGV, this thesis can address the influence of diagnosis of LGV and HIV care and the role of behaviour in those who had LGV (chapter 2 and three), sexual network position is indicated in chapter 3 with an analysis of repeat LGV infection. The role of social context and seroadaptive behaviours in MSM communities in general is explored in chapter 4 and the last chapter explores the potential pathways in which the two infections interact by using a mathematical model of LGV and HIV (chapter 5).

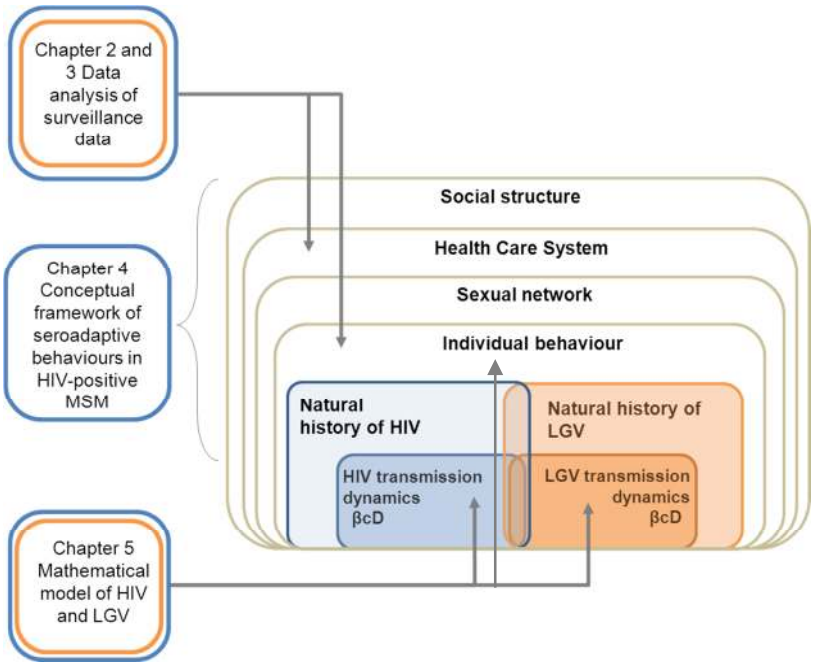
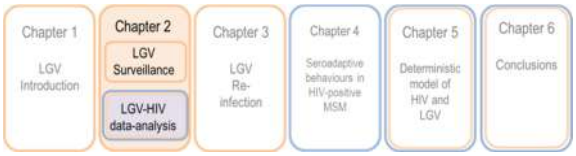


Figure 4. Schematic illustration of factors contributing to the transmission of LGV and HIV in context of this thesis.

Chapter 2

From re-emergence to surveillance:

Data-Analysis of LGV Enhanced Surveillance



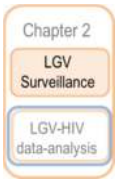
2.1 Summary

Surveillance of STIs is operated mainly through reporting from genitourinary medicine clinics in the United Kingdom. Given its rarity, LGV reporting was previously grouped together with chancroid and donovanosis in the surveillance before its emergence. LGV-specific surveillance was introduced in response to increasing awareness of LGV re-emergence nationally and internationally. LGV surveillance is based on laboratory surveillance at the national STI reference laboratory in Colindale, PHE and it was complemented by LGV Enhanced Surveillance data from the diagnosing clinics which was in place during 2004-2010.

The chapter is divided into two parts. First, I describe the surveillance systems in place for LGV in the United Kingdom, the clinical samples which are tested for LGV, followed by description of the data cleaning for LGV Enhanced Surveillance. Second, I use the dataset to explore factors associated with HIV co-infection in LGV patients, which has been one of the most prominent features of LGV re-emergence. The main aim of this chapter is to compare HIV-diagnosed and HIV-negative/unknown LGV patients to explore whether there are differences in the behavioural and clinical attributes between LGV episodes of different HIV statuses.

I performed a *c*ross-sectional analysis of all cases of LGV in MSM stratified by their HIV status as reported through the LGV Enhanced Surveillance system in the United Kingdom. The data were analysed using logistic regression with generalised estimating equations to control for non-independence of observations due to repeat infections. Three multivariable models were constructed: clinical and behavioural variables were first analysed separately and then a combined model was built to combine explanatory variables associated with HIV status in the dataset.

I show that HIV-positive LGV-infected MSM were more likely to report unprotected receptive anal intercourse compared to LGV-infected MSM with negative/unknown HIV-status. However the data also suggest a diagnostic bias in favour of HIV-positive men who present with a shorter duration of symptoms than men with negative/unknown HIV status. It is therefore possible that HIV-negative men are under-represented in LGV surveillance.



2.2 Surveillance of infectious diseases

Control of infectious diseases requires effective surveillance, which is defined as *"the continuing scrutiny of all aspects of the occurrence and spread of a disease through the systematic collection, collation and analysis of data and the prompt dissemination of the resulting information to those who need to know so that action can result"* (p. 271 in Hawker et al. 2008). Effective surveillance systems generate data which can be used to measure the burden of disease, trends over time and detect emerging health threats, which can guide timely action. Surveillance systems can also identify key populations at heightened risk of infection, inform the allocation of resources, and monitor changes in health practices and the effects of these changes. In addition surveillance can offer a means to describing the clinical course of disease and provides a foundation for epidemiological research (German, Lee, Horan, *et al.*, 2001).

Surveillance is part of prevention efforts which are interventions taking place between health professionals (as well as voluntary and educational sectors) and the public. Prevention efforts are formally divided into primary prevention (health promotion to prevent the illness from occurring through education and changes in society's infrastructure), secondary prevention (detection and treatment, as well as sexual partner management for STIs) and tertiary prevention (management of chronic disease, such as HIV and other long-term conditions) (p. 286 in Pencheon et al. 2008).

Different data sources can be used for surveillance and are discussed by Hawker *et al.* (2008). A case definition including microbiological and clinical criteria forms the basis for surveillance. The type of surveillance system most appropriate for a given infection depends on the natural history of the disease and whether it results in a treatable, acute or chronic condition. Serological surveys give an estimate of the prevalence of infection in the population; and disease can be measured via primary care or hospital information systems whilst others are diagnosed by laboratories. Mortality data on infectious diseases is of limited use in high-income countries with the exception of AIDS, influenza and tuberculosis. Inclusion of particular infections in the statutory notifications systems vary by country. In United Kingdom examples of notifiable infections include food

poisoning, tuberculosis, tropical diseases such as malaria and yellow fever and childhood diseases where there is vaccination such as measles, mumps and rubella.

Detection and treatment are important tools for infection control and, when successful, the prevalence and incidence of infection begin to decline as a result of shorter duration of infection and infectiousness. STI pathogens are able to persist among subpopulations where certain conditions sustain transmission, such as a sufficiently high partner change rate to keep the basic reproductive number[7] (R_0) above 1 even in the presence of external control measures. Alternatively, the pathogen can survive among populations with restricted access to healthcare such as among adolescents and marginalised populations (Wasserheit & Aral, 1996).

Alongside pathogen biology, transmission dynamics are influenced by the socio-cultural context of the host population, which can influence the reproductive number directly through partner change rate, and transmission probability (affected by sexual practices, condom use, circumcision status and host-specific factors such as immunological status). The transmission dynamics are indirectly affected through the socio-economic and geopolitical context in which the host lives such as prevalence of infection in the partner pool, geographical location, globalisation, education, marginality, healthcare services, and cultural and demographic factors (Wasserheit & Aral, 1996; Aral, 2002; Poundstone, Strathdee & Celentano, 2004). These complexities make analysing the impact of sociological factors to the spread of STIs difficult. According to Blanchard (2002) routine surveillance operates from the *pathogen perspective* by examining the time trend of incidence and prevalence of the disease, and lacks the *population perspective* of how the course of an epidemic varies within a population.

2.2.1 Surveillance of sexually transmitted infections in England and Wales

No STI is statutorily notifiable in England and Wales. For HIV there is a variety of data sources for monitoring HIV prevalence, new infections, and disease progression as well as reports of HIV diagnoses in specific groups

[7] Basic reproductive number is defined as "*the average number of secondary cases arising from an average primary case in an entirely susceptible population*" (p.20 in Keeling and Rohani 2008). In its simplest format R_0 for an STI in a homogenously mixing population is determined by its transmission probability, duration of infection and partner change rate.

(Hawker, Begg, Blair, *et al.*, 2008). Reporting of STI diagnoses made in all genitourinary medicine (GUM) clinics in the UK has been mandatory since the network of open access clinics was established in 1917, and until 2009, notification was through completed KC60 (KC denotes Korner code) returns. KC60 returns give aggregate data on the number of diagnosed episodes by STI type, gender and sexuality[8] and are used to assess changes in number of STI diagnoses.

In 2009 a new disaggregated surveillance system was introduced in England: The Genitourinary Medicine Clinic Activity Dataset (GUMCAD) was established in 2009 to replace KC60. GUMCAD electronically collects data on individual-level episodes of STIs (including LGV) and a basic demographic description of the patient. Individual-level data allows, among others, removal of duplicate notifications, and allows for episode- and patient-level analysis, including identification of co-infection and repeat infections. GUMCAD2 has been implemented for non-GUM clinic setting (Public Health England, 2011; Dr Gwenda Hughes, PHE, personal communication).

Enhanced surveillance systems are established for infections which are considered of particular importance for public health, such as for outbreak investigations and for infections where there is limited understanding of the epidemiology of the disease (Pencheon, Guest, Melzer, *et al.*, 2008; Hawker, Begg, Blair, *et al.*, 2008). National Syphilis Enhanced Surveillance (NESS) was in operation from 2002 to 2010 and the Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP) has been on-going from 2000 onwards (Public Health England, 2010a; Jebbari, n.d.). National enhanced surveillance with a shorter duration (from September to December in 2011) was established for shigella in response to a UK-acquired shigella outbreak in MSM (Borg, Modi, Tostmann, *et al.*, 2012). Enhanced surveillance describes the individuals affected and the social context in which the transmission occurs, but this often happens at the expense of smaller sample size as the

[8] The lowest level of aggregation was for a clinic (with unknown catchment area) for a quarter of a year. This type of data was limited by the few variables available to describe the individuals affected, but the data were compensated by good coverage as the vast majority of STI diagnoses in the country were done in GUM clinics or related laboratories and the surveillance had been continuous over a long period giving reliable trend data. KC60 surveillance was supported by voluntary laboratory surveillance, which helps to estimate the proportion of STIs diagnosed outside GUM setting (Hughes, Paine & Thomas, 2001; Ihekweazu, Maxwell, Organ, *et al.*, 2007). However for outbreaks KC60 was too slow and it lacked geographical markers as well as social and behavioural data.

burden of more detailed data collection is laid on the healthcare system (Aral *et al.,* 2002). However the more detailed information collected of subset of the individuals can provide a needed population perspective (Blanchard, 2002).

2.2.2 Development of LGV surveillance in United Kingdom

LGV was considered to be a rare tropical disease, and it was previously grouped together with chancroid and donovanosis in KC60 returns. This prevents estimates of baseline level of LGV occurrence prior to its re-emergence. Moreover, prior to 2004 the diagnosis was not standardised and was often based on clinical presentation rather than established laboratory protocol.

Countries which have a case definition for LGV require identifying the biovar (serovars L1-L3) for case confirmation (Timen, Hulscher, Vos, *et al.*, 2008). As the standard chlamydia test does not differentiate between LGV and non-LGV serotypes, most laboratories rely on characterising the *ompA* gene after the sample is confirmed positive for *C. trachomatis* (Morre, Ouburg, van Agtmael, *et al.*, 2008). *ompA* codes for the major outer membrane protein (MOMP) which is the primary LGV surface antigen (Harris, Clarke, Seth-Smith, *et al.*, 2012).

After the international alert from European Network for Surveillance of STIs in 2004 as described in chapter 1, new protocols for the detection of LGV were established in the United Kingdom (von Holsterin, Fenton & Ison, 2004). For an effective surveillance system there first needed to be a diagnostic capacity for LGV – the previous syndromic approach was insufficient – and the confirmed LGV episodes needed to be linked to patient data. In the UK[9], LGV diagnosis became centralised in Sexually Transmitted Bacteria Reference Laboratory (STBRL) due to the diagnostic methods required for detection of LGV serovars.

[9] From August 2006, Scottish specimens were referred to the Scottish Bacterial Sexually Transmitted Infections Reference Laboratory. These results are still reported to Colindale to give UK wide surveillance data (Dr Gwenda Hughes, PHE, personal communication).

STBRL tests specimens for LGV from patients with symptoms and/or who are contacts of an LGV positive patient (Public Health England, 2010b). In STBRL Referral guidance (Sexually Transmitted Bacteria Reference Laboratory, n.d.) samples positive for chlamydia from patients with clinical symptoms suggestive of LGV or sexual partners of LGV infected are invited to be sent for LGV testing[10]. Samples arriving to STBRL for LGV testing are first re-screened for *C. trachomatis* (by real-time PCR and independent primers (Chen, Chi, Alexander, *et al.*, 2007)). A small proportion (5-10%) of the samples degrade during transit, and some test negative for *C. trachomatis* due to differences in sensitivity of different testing methods (Sarah Alexander, STBRL , personal communication).

A voluntary enhanced surveillance system was introduced for LGV in 2004 after the initial outbreak alert had identified the first cases. It was discontinued at the end of 2010. The LGV Enhanced Surveillance used a paper-based form consisting of 24 main questions and associated sub-questions (form in Appendix for Chapter 2) which was filled by the clinician retrospectively (based on clinical notes and/or following consultation with the patient) after the patient had been confirmed to have LGV[11].

In the preceding sections, I described the importance of surveillance to understanding the epidemiology of diseases to facilitate effective disease prevention. The first two sections also provided the historical and clinical context of enhanced surveillance of LGV in the UK. Prior to any data analysis, it is first important to describe

---

[10] STBRL recommend the following in their LGV laboratory sample referral form "*The LGV service is free of charge for specimens which have been confirmed as Chlamydia trachomatis positive at the local laboratory using a nucleic acid amplification test (NAAT) and have been sourced from either a symptomatic patient or a direct sexual contact. Ideally STBRL will accept rectal swabs from patients with proctitis or urethral swabs, urine or lymph node aspirates from patients with inguinal lymphadenopathy. Either the residual processed NAAT specimen or a dry unprocessed specimen will be accepted. In rarer instances STBRL will also accept ulcer swabs, lymph node biopsies, rectal biopsies, and cervical swabs, if there is clinical suspicion of LGV. Please contact the STBRL before sending*".

[11] In relation to other European countries the UK has done well with laboratory surveillance covering the whole country, enhanced surveillance system in place between 2004-2010 and guidelines on LGV management were put together (Clinical Effectiveness Group of the British Association for Sexual & HIV, n.d.) as well as PHE giving concise information on its website. In a cross-sectional survey of LGV surveillance performed in Europe, 7/11 of countries who had reported cases had established an enhanced surveillance system for LGV based on voluntary reporting by early 2006 (Timen, Hulscher, Vos, *et al.*, 2008). They also commented on the variability in surveillance measures taken across the countries.

the nature of the raw data and a detailed account of the steps in the data cleaning process. In the following section I describe two principal sources of data: the LGV laboratory surveillance data (in section 2.3.1.1), and the LGV Enhanced surveillance data (in section 2.3.1.2). A summary of the data cleaning process to prepare the data for the subsequent analyses is provided in section 2.3.2 with more detailed information provided in appendix for chapter 2, Table S1.

2.3 LGV surveillance data description

2.3.1 Data sources

2.3.1.1 LGV laboratory surveillance

A record of samples that were tested for LGV was obtained from STBRL , PHE. This was used to describe the testing patterns by infection site and sex of the patient. It consists of 11,196 *C. trachomatis* positive samples that were tested for LGV in STBRL from 2004 until the end of 2010. Table 1 describes the sample stratified by sex and site of infection. The vast majority of samples tested for LGV in men are from rectal swabs while 6.7% of samples are taken from non-rectal sites; the rest have an unknown sample site.

Overall 15.5% (1,417/9,138) of rectal swabs from men were found to have LGV with urethral swab, urine and throat swab LGV positivity was 7.1% (11/156), 2.9% (4/136) and 2.7% (4/151), respectively. Only 0.8% of samples from women (4/531) tested positive for LGV, all from rectal swabs.

Table 1. Chlamydia positive samples tested for LGV at STBRL, stratified by sex and site of infection.

MALE

	Total (% of total)		LGV+		LGV-		Indeterminate	
	n	**%**	n	%	n	%	n	%
Rectal Swab	**9,138**	**91.1%**	1,417	15.5%	7,718	84.5%	3	0.0%
Urethral Swab	**156**	**1.6%**	11	7.1%	145	92.9%	0	0.0%
Urine	**136**	**1.4%**	4	2.9%	130	95.6%	2	1.5%
Throat swab	**151**	**1.5%**	4	2.6%	147	97.4%	0	0.0%
Other	**230**	**2.3%**	46	20.0%	184	80.0%	0	0.0%
Unknown	**224**	**2.2%**	41	18.3%	183	81.7%	0	0.0%
Total	**10,035**	**100.0%**	**1,523**	**15.2%**	**8,507**	**84.8%**	**5**	**0.0%**

FEMALE

	Total (% of total)		LGV+		LGV-		Indeterminate	
	n	**%**	n	%	n	%	n	%
Rectal Swab	**451**	**84.9%**	4	0.9%	447	99.1%	0	0.0%
Other sites	**80**	**15.1%**	0	0.0%	80	100.0%	0	0.0%
Total	**531**	**100.0%**	**4**	**0.8%**	**527**	**99.2%**	**0**	**0.0%**

630/11,196 (5.6%) with unknown sex are not presented in these tables

2.3.1.2 LGV Enhanced Surveillance

Another dataset was obtained from STI Section at the Centre for Infectious Disease Surveillance and Control in Colindale, PHE (previously known as the HIV and STI Department of HPA) of LGV confirmed episodes which had a completed LGV Enhanced Surveillance form. These data had been collected and maintained by PHE who performed preliminary data linking and cleaning. The dataset was in Excel Worksheet (1997-2003) format which was converted to Stata database using Stat transfer programme. The dataset includes 1370 LGV episodes from 2003 (with the first episode recorded retrospectively in 2004) to the end of 2010. Laboratory confirmed samples (LGV laboratory data in Table 1) and LGV episodes in the Enhanced Surveillance dataset cannot be directly compared as multiple laboratory samples may in some cases come from the same patient but the coverage of the enhanced surveillance is approximately 87%[12].

2.3.2 Data cleaning

The LGV laboratory data was used for description purposes only, and this section concerns cleaning of data from the LGV Enhanced Surveillance. LGV Enhanced Surveillance form is presented in Appendix for Chapter 2. The majority of responses to questions on the paper-based LGV Enhanced Surveillance form were pre-defined multiple-choice options requiring a "tick" if applicable to the patient, as well as an "other" text field allowing

[12] PHE estimate that LGV Enhanced Surveillance data was available for 87% (1,370/1,581) of LGV cases (after de-duplication) (Dr Gwenda Hughes, PHE, personal communication).

further specification of the response if none of the pre-defined options applied to them. Most questions dealt with the existence of conditions and prior behaviour, and few questions queried if the patient did not have a specific condition. This resulted in a database where the item response varies between variables, and the text fields add to the variability in information. An extensive data cleaning was performed which is detailed in Appendix for chapter 2, table 1 S. Many of the Enhanced Surveillance Form questions are structured in similar way: under a question several options were displayed of which one or more may have been selected. A standardised procedure for data cleaning was followed:

- For a particular question, information for the option was assumed missing if none of the options or the text field was filled;
- The option was considered negative if it was not selected but another point under the same question was selected instead.
- A dummy variable was created for the unrecorded category (including those who had ticked "unknown"). This resulted in the unknown category to be identical between the options under the same question as all fields had to be missing for the question to be considered unrecorded.

The observations in the dataset are referred to as (LGV) episodes. Data cleaning was performed in Stata SE/11. In Table 2 the variables measured in LGV Enhanced Surveillance are summarised. In the same Appendix Table S 2 present a descriptive tabulation of the variables in the dataset.

Table 2. Summary of variables in the LGV Enhanced Surveillance

Demographic variables	Clinical variables	Behavioural variables

• Gender
• Sexuality
• Age
• Ethnicity
• City of clinic
	• Date of symptoms onset
	• Date of presentation
	• Reason(s) for attending
	• Site of infection
	• Types of symptoms
	  ○ Genital
	  ○ Rectal
	  ○ Systemic
	• Treatment
	• Concurrent STIs
	  ○ Gonorrhoea
	  ○ Chlamydia
	  ○ Non-specific urethritis (NSU)
	  ○ Syphilis
	  ○ Genital warts
	  ○ Genital herpes
	  ○ Hepatitis B
	• Hepatitis C status
	• HIV status
	  ○ Date of HIV diagnosis
	  ○ CD4 count
	  ○ HAART
		• Probable country of acquisition
		• Location for meeting new partners
		"Sex-on-premises" venues:
		  ○ Backroom
		  ○ Cruising ground
		  ○ Sauna
		  ○ Sex party
		Others (more mainstream venues):
		  ○ Bar/club
		  ○ Internet
		• Number of sex partners
		• Types of sex LGV patients have had in the past three months
		  ○ Receptive anal intercourse (RAI)
		  ○ Insertive anal intercourse (IAI)
		  ○ Oral sex
		  ○ Sharing sex toys
		  ○ Fisting

Response rates between variables varied (see table S 2 in Appendix for chapter 2). In general, demographic and clinical questions had a higher response rate. The exception was for specialised diagnostic tests such as for hepatitis C with the PCR test missing for over half (54.6%), and the antibody (Ab) test having 19.8% of responses in the unknown category. For behavioural variables more information was missing, for example, over 50% had no information on locations for meeting new partners. For sexual practices, questions on oral and anal sex were relatively well completed (with maximum of 21.0% unknown for insertive anal intercourse), but more information was unrecorded for the more esoteric options such as fisting and sharing sex toys, with almost half of the responses missing.

For later analyses several variables were combined for ease of interpretation and due to underlying correlation[13]: When the correlation was high, the categorical variables were combined where possible, or if no meaningful combination was practical, one of the variables was left from the multivariable model. For concurrent STIs there was a strong correlation (>0.5) and a variable "any concurrent STI" was created (including gonorrhoea, non-specific urethritis (NSU), syphilis, genital warts, genital herpes and hepatitis B). A very strong correlation (>0.9) was present among variables describing locations for meeting new sexual partners, mainly caused by the common unrecorded category for these variables. Therefore two collapsed variables "met partners in any of these locations" and "met partners in sex-on-premises locations" were created. For sexual practices there was a strong (>0.8) correlation between insertive and receptive oral sex variables, where most who reported unprotected oral sex in one, also reported unprotected oral sex in the other A similar situation occurred for insertive and receptive fisting (>0.9 correlation, due to most episodes not having reported either or both having the variables unrecorded). Collapsed variables were created for both of these sexual practices. Sharing sex toys and "any fisting" were also correlated (>0.8), and based on cross-tabulation the correlation seemed to be similarly concentrated in categories "None reported" and "unknown/missing". There was no meaningful way to combine these two variables, so fisting was selected as the more relevant (based on literature) variable to include in later analyses. In preliminary analysis, any genital symptoms and any rectal symptoms were identified as potentially important explanatory variables and these were collapsed into a composite variable representing genital-and-rectal-symptoms.

2.3.3 Identification of repeat infections

For those with more than one episode of LGV in the dataset, I estimated the episode number based on a separate data source provided by PHE. This gave a laboratory ID, Enhanced Surveillance ID – if the episode had an enhanced surveillance form – and a repeat linkage identifier through which it was possible to link episodes in the LGV Enhanced Surveillance dataset. Using the date of clinic presentation and the two ID numbers (both

[13] Correlation (covariance) measures the strength of association between two variables; the correlation coefficient can take values from -1 to 1 with 0 for no association between the variables. When two variables have a high correlation they are said to be collinear. This can be problematic in multivariable models as they can cancel the association they have independently with the outcome when both are included in the regression model leading to an incorrect interpretation that neither is associated with the dependent variable (Kirkwood & Sterne, 2008).

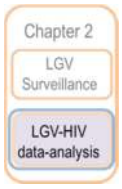
are consecutive positive integers), I estimated whether the episode was the patient's first, second or third recorded episode. In cases where two episodes within an individual occurred less than three months apart, the second was excluded as a possible duplicate notification or treatment failure, which has been indicated as a probable source of re-infection in the first weeks after treatment (Whittington, Kent, Kissinger, *et al.*, 2001). This information was used to create an identifier that linked repeat infections in an individual.

2.3.4 Definition of HIV status during LGV episode

As the following data analysis compares HIV-diagnosed and HIV-negative/unknown LGV patients to each other I will describe how the HIV status was defined in this section. The dependent variable in the statistical analysis was HIV status, and in this case I was using HIV status as an indicator of sub-population membership with the hypothesis that, as a population, HIV-positive MSM are different to MSM who are HIV-negative (or do not know their HIV status). I hypothesise that HIV-positive MSM form a community of MSM whose behaviour and social and sexual networks are different to HIV-negative men (Williamson, Dodds, Mercey, *et al.*, 2008). Furthermore this affects the interpretation of the results as the outcome is being HIV-positive (rather than acquiring HIV).

Items 14-18 of the Enhanced Surveillance form were used to determine whether the LGV episode occurred in a known HIV-positive individual or in an individual who is not known to be HIV-positive. How each question was treated in the interpretation of the patient's HIV status is stated in Table S1 (Appendix for chapter 2) next to the items 14-18. This leaves a comparison group of presumed HIV-negative individuals. However no question identifies whether the HIV-negative LGV episodes had an HIV test at LGV diagnosis. Even though the majority of those not identified as HIV-positive were consistently marked as not having HIV, and HIV testing is likely to happen for a patient who presented with LGV, the comparison group is more accurately defined as HIV-negative/unknown, resulting in potential misclassification of the main outcome for some episodes. However, uptake of HIV-testing among eligible MSM at GUM clinics is around 90-93% (of those who were offered an HIV test in 2009-2011), which would suggest that the majority of LGV cases who are not known to be positive have tested HIV-negative (Table 4 in Public Health England, 2010d).

This forms the basis of the population used for data analysis in the next part of the chapter, which explores the differences between HIV-positive and HIV-negative/unknown LGV patients, and the chapter that follows which looks at repeat infection in more detail. LGV enhanced surveillance offers a cross-sectional outlook into the episodes at the clinic presentation and factors preceding this presentation. To our knowledge the LGV Enhanced Surveillance is the largest LGV related dataset collected to date and the dataset can offer a meaningful insight into those who were diagnosed with LGV.



2.4 Data-analysis of HIV status in LGV Enhanced Surveillance dataset

HIV-infected sexually active MSM have been a predominant feature of LGV emergence across affected countries. As discussed in the introductory chapter it is unclear whether the association between HIV and LGV is due to common risk factors, distinct and overlapping sexual networks, biological synergy or whether it could be in part an artefact of diagnostic systems (Ward, Martin, Macdonald, *et al.*, 2007).

If risk behaviour and HIV-status are linked we would expect HIV-negative individuals to have less risky sexual behaviours, and perhaps less likely to be transmitting LGV. (On the other hand, since HIV acquisition is determined by a probabilistic process across an individual's sex acts, there is likely a small proportion of people with high-risk behaviours who will be HIV-negative at a particular point in time.) Previous observational studies have noted an increase in reporting of unprotected anal intercourse in both HIV-positive and –negative MSM (Lattimore, Thornton, Delpech, *et al.*, 2011a), and HIV-positive men, on average, report more risk behaviour (in terms of partner numbers and unprotected anal sex) compared to HIV-negative men (Williamson, Dodds, Mercey, *et al.*, 2008). Dougan *et al.* (2007) reviewed the literature on HIV prevalence in MSM with STIs and found HIV prevalence to be higher in MSM with STIs than MSM from community samples. They identify gonorrhoea, syphilis, hepatitis C and LGV as recently re-emerged STIs in MSM, but only hepatitis C has been more closely associated with HIV than LGV with almost all cases seen in HIV-positive men (Danta, Brown, Bhagani, *et al.*, 2007).

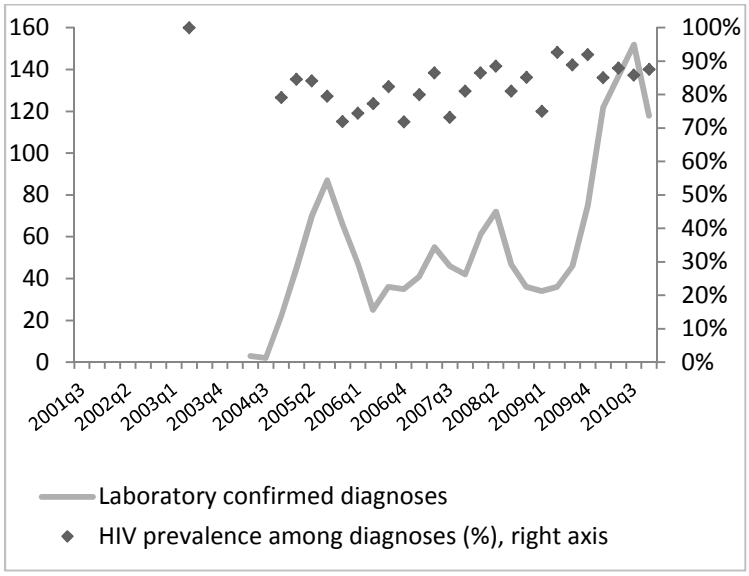


Figure 1. Number of laboratory confirmed diagnoses by quartile of year, and HIV prevalence in LGV Enhanced Surveillance data, presented on the right axis.

Figure 1 compares the number of confirmed LGV episodes against the HIV prevalence in LGV episodes as observed in the LGV Enhanced Surveillance data. We might expect the infection to start from those with the highest risk for acquisition, here HIV-positive MSM, followed by a decline in HIV prevalence as the epidemic matures and infections spreads outside the "core groups". However, the surveillance data shows a steady level of HIV prevalence in LGV cases throughout the surveillance period.

This study is observational in nature and we cannot directly observe biological or behavioural properties in relation to HIV-positivity. Neither can we directly examine network structure given we have no information beyond the individual and also since everyone in the data has LGV we cannot explore differences in acquisition risk by HIV status. From this design we can draw inferences on clinical and behavioural attributes on an individual level. If there were clear differences in the clinical presentation or behaviours reported, this analysis may give us an indication of the nature of the association. This gives a more comprehensive picture of LGV cases – as long as they have been captured by the surveillance system – and for the factors that have been measured for. The design of the data and methods used allows exploration of the objectives listed below.

2.5 Objectives

The main aim of this analysis is to identify key differences between HIV-positive and HIV-negative subpopulations with LGV, and explore potential reasons for the high level of HIV in LGV cases. This is achieved by three specific objectives:

- To identify potential differences in the clinical presentation between those with and without HIV through the construction of a multivariable model of clinical factors.
- To identify potential differences in sexual risk behaviour and indicators of sexual network membership prior to LGV acquisition between those with and without HIV through the construction of a multivariable model of behavioural factors.
- To identify the overall differences between HIV-positive and HIV-negative sub-populations who acquire LGV through the construction of a statistical model including both clinical and behavioural factors.

2.6 Methods

2.6.1 Statistical analysis

Logistic regression is a common choice of statistical model when the outcome is binary as described in Hosmer and Lemeshow (2000). Due to the binary format of the outcome logistic regression differs from linear regression in a few important ways: the conditional mean of the outcome (given the exposure variables) has a logit transformation so it lies between 0 and 1, and binomial distribution defines the distribution of errors and is the basis for the statistical analysis. Subsequently to interpret the estimated coefficient it needs to be transformed by taking the exponential of the coefficient resulting in odds ratio (OR) (Hosmer & Lemeshow 2000, p. 49-50).

In standard logistic regression independence of observations is assumed. If the independence assumption is violated the standard errors and resulting confidence intervals can become biased. In this dataset episodes belonging to the same individual are more likely to be similar than episodes belonging to different individuals, and therefore analysis was performed at individual-level allowing clustering for repeat infections.

I performed statistical analysis using a logit model and generalised estimating equations (GEE) and robust standard errors; the method was described in a paper by Liang and Zeger (1986). GEE, is a population-average model, which takes into consideration correlation within clusters, but assumes no between cluster correlation (thus making it different from random-effects model) (Hu, Goldberg, Hedeker, *et al.*, 1998). The within cluster dependency is accounted for in the robust standard errors which correct for the population variance according to the clusters (Hu et al. 1998) which results in different standard errors compared to logistic regression where an independent correlation structure is assumed when calculating the standard error (Kirkwood & Sterne, 2008). An exchangeable correlation structure was selected as recommended by Agresti (2002, p. 468) who advises using an exchangeable correlation structure if large differences in the correlations are not expected; also Hosmer and Lemeshow (2000, p. 312-314) note that exchangeable correlation assumption works well under most circumstances. This approach recognises the within cluster dependence but uses only one additional parameter and it assumes uniform correlations within clusters across time (Hu, Goldberg, Hedeker, *et al.*, 1998). In summary the larger the cluster the less relative weight an observation within the cluster gets, and this helps to control the overall influence of individuals who acquire LGV more than once. I performed the analysis using Stata/SE 11.2 (the command for the GEE logit model used was `xtlogit depvar [indepvars], pa corr(exch) i(id) robust or`). Stata uses the Huber/White estimator of variance and the resulting standard errors are labelled as semi-robust instead of robust.

2.6.1.1 Multivariable model building strategy

The aim with the statistical models in this study was to develop an explanatory model of those variables that have an influence on the dependent variable. Explanatory models aim to understand why empirical phenomena occur. According to Kirkwood and Sterne (2008) the focus in explanatory models is not explicitly "*on identifying which confounders to include for a particular risk factor, nor is it on identifying any combination of exposures that works, as in the prediction scenario*" (p 342); for an explanatory model they recommend against formal stepwise methods and recommend a conceptual framework instead, with the selection of variables according to their predetermined significance to the outcome.

The fewer variables there are in the statistical model, the more numerically stable it is likely to be. In addition this makes the model easier to generalise (Hosmer & Lemeshow, 2000). The general guidance is to have

approximately ten times the number of observations as there are indicator variables (Kirkwood & Sterne, 2008). In this analysis another limiting factor is the smaller number of observations in the control (HIV-negative/unknown) group compared to the cases.

To take into consideration the above factors two preliminary multivariable models were constructed: one for clinical factors and another for behavioural factors. Of these a final multivariable model was produced of variables that were considered to be of interest for the HIV-LGV association or to be potential confounders for this association according to *a priori* hypotheses and objectives. I deduced that for behavioural variables more proximate determinants for LGV acquisition are likely to be a better measurement of differences in sexual risk behaviour than more distal determinants. I also thought that differences in clinical pathways might be evident in the dataset (such as presence of other STIs that may increase suspicion for LGV, or differences in type of symptoms reported). Secondary selection criteria were based on statistical "importance" where a p-value 0.2 was used as a cut-off point for statistical association of interest.

2.6.1.2 Variables

Variables describing age, ethnicity, sexuality, probable acquisition country, number of sex partners, meeting new partners at sex-on-premises venues and sexual practices preceding the LGV episode were analysed as behavioural variables (ethnicity, age and sexuality were categorised under behavioural variables as I assessed these factors to be more related to behaviour than to clinical presentation). For this analysis hepatitis C PCR results were selected instead of hepatitis C antibody (Ab) results as a positive PCR test infers acute hepatitis C infection[14]. I selected meeting new partners in sex-on-premises venues (backroom, cruising ground, sauna and sex party) for this analysis instead of "any venue" (which also includes the Internet and bar/club) as it offers a more meaningful inference of risk behaviour and potential network membership preceding the LGV episode.

Age, duration of symptoms, number of sex partners, presentation day to the clinic and days between presenting to the clinic and filling in the surveillance form ("form delay") were originally continuous variables. Categorical variables were created during data cleaning depending on the distribution and interpretation of

[14] This gives an approximation in the absence of previous hepatitis C antibody or RNA results. Chronic and acute hepatitis C is defined and reviewed by Bradshaw *et al.* (2013).

each of these variables. The Wald test was used to test whether the continuous or categorical variable would be more appropriate given the dataset[15].

For presentation date (to the clinic), categorical and continuous variables gave equally good fits to the data, and I decided to select a categorical variable (episode presented in 2010 compared to before 2010). This was done to reflect the sudden increase in case-numbers during 2010 and potential impact this might have had on the case profile. Age and number of sex partners were not significantly different between the categorical and continuous forms of the variable, and these were entered into the model as continuous variables. Duration of symptoms (prior to presenting to the clinic) in its categorical form (based on mode of 7 days) gave a better fit to the data than in continuous form, and the categorical variable was entered into the model (categorised as week or less, more than a week or unknown).

Duration of form delay (number of days between presentation to the clinic and form completion) was not thought to be a variable of interest in itself. Form delay was considered to be potentially affecting how well the exposure variables were measured[16] and form delay was entered in all multivariable models as a continuous variable.

When the correlation was high the categorical variables were combined where possible (as described in the previous data cleaning section), or if no meaningful combination was practical, one of the variables was excluded from the multivariable model. This was considered justifiable, since when the correlation between the variables is high, the variables can be assumed to be measuring a similar phenomenon and are therefore collinear with respect to the outcome (Kirkwood and Sterne, p337-9).

---

[15] I used `testparm varlist, equal` command in Stata to compare whether the coefficients of the variables were equal to each other (in a previously fitted model). When the difference was statistically significant I used the variable which gave a better fit; where there was no statistically significant difference the continuous variable was favoured as it uses fewer degrees of freedom compared to the categorical variable. The likelihood ratio test, commonly used for logistic regression diagnostics, was not appropriate here since GEE is not based on likelihood theory like the standard logistic regression (Hosmer & Lemeshow, 2000).

[16] Completeness of form is likely to decline the longer it takes to fill it (see figure S 2 in Appendix for Chapter 2 describing the delay in data collection with a median time of 98 days between presentation to the clinic and filling in the enhanced surveillance form).

2.7 Results

The LGV Enhanced Surveillance dataset is composed of 1,370 LGV confirmed episodes. Figure 2 presents episodes that were excluded: 3 episodes occurred in females, 2 in those of unknown sex, 5 in heterosexual men, and 10 episodes occurred in men with unknown sexuality. Eight episodes occurred under 3 months after the previous episode in the same individual, and were excluded.

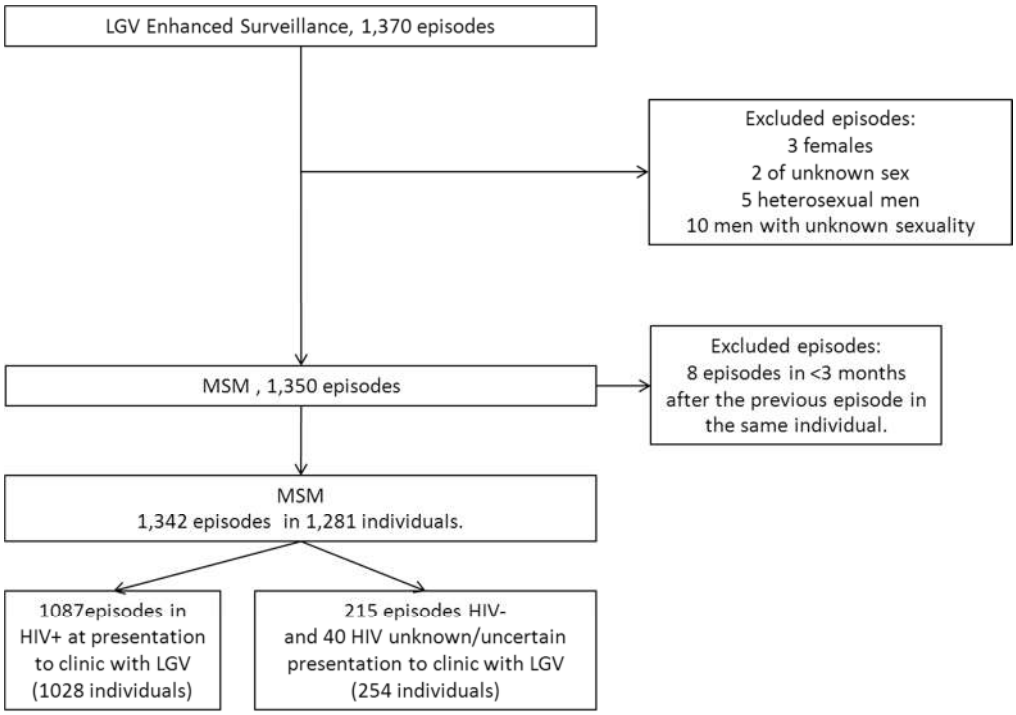


Figure 2. Excluded and included episodes in the LGV Enhanced Surveillance dataset

This resulted in a dataset of MSM with information of their HIV status at presentation to clinic (including those who would have been diagnosed with HIV during LGV episode), and is comprised of 1087 HIV-positive episodes in 1028 individuals and 254 HIV-negative or status-unknown individuals. (In 40/255 episodes (15.7%) the HIV status was unknown or uncertain (inconsistently reported or missing) at the time of their LGV diagnosis and these were included in the "HIV-negative/unknown group".)

Of 1,342 episodes, 65 (4.8%) were identified as repeat infections and all apart from one occurred in the known HIV-positive group. The distribution of episodes along with the estimated episode number is presented in Table 4. In one patient HIV was diagnosed between first and second LGV episode.

Univariate analyses along with distribution of the variables that were then collapsed are presented in the Appendix for chapter 2 in Table S 3 whilst univariate and multivariable models are presented in this chapter in Table 4-Table 6. Correlation coefficients between variables chosen for the final models are presented in Appendix for chapter 2 Table S 7.

2.7.1 Clinical presentation

In the clinical correlates for HIV status in univariate analysis, presented in Table 4, having been seen in a clinic in London was positively associated with being HIV-positive (OR 1.6, 95% CI 1.2, 2.1), as was having a positive hepatitis C (PCR) result (OR 5.9, 95% CI 2.0, 17.5). However there were few events among HIV-negative/unknown (n=3), therefore the effect size should be interpreted with caution and is only indicative of a qualitative trend.

The following variables had a significant positive or negative association with HIV-positive status at the 0.05 level: more than a week's duration of symptoms (OR 0.6, 95% CI 0.5, 0.8), presenting to the clinic via referral (OR 0.5, 95% CI 0.2, 0.8), presenting to the clinic in 2010 (OR 1.6, 95% CI 1.2, 2.2), reporting only genital symptoms (OR 0.32, 95% CI 0.14, 0.71) and having infection in the genital region (OR 0.30 95% CI 0.10, 0.90). There was a negative association between being HIV-positive and having unrecorded information for three variables: hepatitis C (PCR) (OR 0.39, 95% CI 0.28, 0.53), site of infection (OR 0.5, 95% CI 0.37, 0.67) and other concurrent STIs (OR 0.59, 95% CI 0.40, 0.88).

I constructed a multivariable model containing all clinical variables that were statistically associated with the outcome at the univariate level or considered a priori to be of potential interest, presented also in Table 4. Interestingly all but one of the recorded re-infections with LGV occurred in the HIV-positive group. This prevented a comparison of recurrent infections, or controlling for the episode number in the multivariable analysis in ways other than allowing for clustering using GEE. Also hepatitis C infection was so strongly correlated with HIV status that I decided to exclude it from multivariable analysis. Among reasons for visiting the clinic, only clinic referral was significantly associated with being HIV-positive; symptoms, contact tracing and routine STI screen were evenly distributed across the outcome (Table 4, as demonstrated by the

descriptive tabulation and univariate results with all p-values above 0.2) and were not included in the multivariable analyses.

In the multivariable model, looking at explanatory variables for being HIV-diagnosed, form delay is controlled for in addition to the variables presented in Table 4. Overall, the multivariable model of clinical variables did not reveal major changes in the association of the variables with the outcome compared to the univariate analysis. Being seen in a clinic in London became a less important explanatory variable for being HIV-positive with a wider confidence interval and reduced adjusted odds ratio (aOR 1.37, 95% CI 0.94, 1.98); whilst having more than a week of symptoms reported maintained a significant negative association (aOR 0.52, 95% CI 0.34, 0.78). Confidence intervals for the type of symptoms and site of infection reported became wider and genital symptoms (aOR 0.46, 95% CI 0.17, 1.22) and genital infection (aOR 0.56, 95% CI 0.15, 2.03) ceased to be statistically significant. Conversely, reporting any systemic symptoms increased in effect size and became statistically significantly associated with being HIV-positive (aOR 1.57, 95% CI 1.04, 2.38). A similar phenomenon occurred with referral as a reason for attending (referral from another clinician to the STI/HIV clinic), which had a negative association with being HIV-positive when controlling for other clinical factors (aOR 0.33, 95% CI 0.13, 0.80). The multivariable model had 969 observations in 924 groups, and had 26 explanatory indicator variables (including categories) resulting in 35.5 groups per variable level (924/26).

Surprisingly, site of infection and reported symptoms were not correlated (correlation 0.014), and to explore the reasons I cross-tabulated these, stratified by HIV status in the Appendix for chapter 2, Table S 4. Overall, for a majority of episodes, an unknown site of infection is recorded: 71.0% (181/255) and 56.5% (614/1087) of episodes have an unknown site of infection for HIV-negative/unknown MSM and HIV-positive MSM, respectively. Reporting of infection site was better in 2010, and to aid interpretation, I have presented the same tabulation restricted to episodes seen in 2010 (Table 3). In 2010, the majority of episodes with a rectal infection site also reported only rectal symptoms, and did not vary by HIV status (68.0% and 69.1%). No one with a genital LGV infection reported rectal symptoms. Though the numbers with a genital site of infection are small, even among rectal infection sites more than 2/3 of symptoms were reported as rectal only. This

suggests that site of infection and site of reported symptoms are correlated but that the high proportion of unknown site of infection prior to 2010 obscures this trend.

Table 3. Cross-tabulation of site of infection and reported symptoms stratified by HIV status. The sample is restricted to LGV episodes in 2010 with most episodes having infection site reported. The percentage value is of the row total.

HIV-negative/unknown (year 2010)

		Reported symptoms					
Site of infection		None	Only genital	Only rectal	Both	One or both unknown	Total
Rectal	n	3	1	34	11	1	50
	%	**6.0**	**2.0**	**68.0**	**22.0**	**2.0**	**100.0**
Genital	n	0	1	0	0	0	1
	%	**0.0**	**100.0**	**0.0**	**0.0**	**0.0**	**100.0**
Both/throat	n	1	0	0	2	0	3
	%	**33.3**	**0.0**	**0.0**	**66.7**	**0.0**	**100.0**
Unknown	n	0	0	3	0	1	4
	%	**0.0**	**0.0**	**75.0**	**0.0**	**25.0**	**100.0**
Total	n	4	2	37	13	2	58
	%	**6.9**	**3.5**	**63.8**	**22.4**	**3.5**	**100.0**

HIV-positive (year 2010)

		Reported symptoms					
Site of infection		None	Only genital	Only rectal	Both	One or both unknown	Total
Rectal	n	21	10	212	48	16	307
	%	**6.8**	**3.3**	**69.1**	**15.6**	**5.2**	**100.0**
Genital	n	1	6	0	0	0	7
	%	**14.3**	**85.7**	**0.0**	**0.0**	**0.0**	**100.0**
Both/throat	n	0	0	2	2	0	4
	%	**0.0**	**0.0**	**50.0**	**50.0**	**0.0**	**100.0**
Unknown	n	2	4	21	3	1	31
	%	**6.5**	**12.9**	**67.7**	**9.7**	**3.2**	**100.0**
Total	n	24	20	235	53	17	349
	%	**6.9**	**5.7**	**67.3**	**15.2**	**4.9**	**100.0**

Table 4. Association between clinical factors and being HIV-positive in MSM with LGV: results of univariate and multivariable analysis

	HIV+ n=1087		HIV-/unknown n=255		Univariate logistic regression (GEE)				Multivariate logistic regression * (GEE)			
	n	%	n	%	OR	1.0	CI	P-value	OR	1.0	CI	P-value
Presentation year												
Before 2010	731	**67.7**	196	**77.2**	1.0				1.0			
During 2010	349	**32.3**	58	**22.8**	1.6	1.2	2.2	0.003	1.0	0.7	1.5	0.989
Episode number												
1st	1023	**94.1**	254	**99.6**	Not included				Not included			
2nd	58	**5.3**	1	**0.4**								
3rd	6	**0.6**	0	**0.0**								
Seen in a clinic in London												
No	316	**29.1**	100	**39.2**	1.0				1.0			
Yes	771	**70.9**	155	**60.8**	1.5	1.2	2.1	0.003	1.4	0.9	2.0	0.099
Duration of symptoms												
Week or less	377	**34.7**	65	**25.5**	1.0				1.0			
More than a week	500	**46.0**	141	**55.3**	0.6	0.5	0.8	0.002	0.5	0.3	0.8	0.002
Unknown	210	**19.3**	49	**19.2**	0.7	0.5	1.1	0.146	0.7	0.4	1.3	0.304
Reasons for attending the clinic												
Symptoms												
No	139	**12.8**	31	**12.2**	1.0				Not included			
Yes	925	**85.1**	218	**85.5**	0.9	0.6	1.4	0.783				
Unknown	23	**2.1**	6	**2.4**	0.8	0.3	2.1	0.710				
Contact tracing												
No	980	**90.2**	229	**89.8**	1.0				Not included			
Yes	84	**7.7**	20	**7.8**	1.0	0.6	1.6	0.927				
Unknown	23	**2.1**	6	**2.4**	0.9	0.4	2.1	0.769				
Routine STI screen												
No	988	**90.9**	232	**91.0**	1.0				Not included			
Yes	76	**7.0**	17	**6.7**	1.1	0.6	1.8	0.849				
Unknown	23	**2.1**	6	**2.4**	0.9	0.4	2.1	0.779				
Referral												
No	1,033	**95.0**	233	**91.4**	1.0				1.0			
Yes	31	**2.9**	16	**6.3**	0.4	0.2	0.8	0.011	0.3	0.1	0.8	0.015
Unknown	23	**2.1**	6	**2.4**	0.8	0.4	2.0	0.713	0.6	0.2	1.6	0.292
Location of symptoms reported												
None	60	**5.5**	11	**4.3**	1.0				1.0			
Only Genital	41	**3.8**	24	**9.4**	0.3	0.1	0.7	0.005	0.5	0.2	1.2	0.116
Only Rectal	758	**69.7**	147	**57.7**	1.0	0.5	1.8	0.880	1.2	0.5	2.8	0.600
Both	172	**15.8**	54	**21.2**	0.6	0.3	1.2	0.140	0.7	0.3	1.7	0.403
Unknown	56	**5.2**	19	**7.5**	0.6	0.3	1.3	0.167	0.9	0.3	2.4	0.810
Site of infection												
Rectal	454	**41.8**	63	**29.3**	1.0				1.0			
Genital	10	**0.9**	5	**2.3**	0.3	0.1	0.9	0.032	0.6	0.2	2.0	0.377
Both or other (throat, n=1)	8	**0.7**	3	**1.2**	0.4	0.1	1.5	0.182	0.6	0.1	2.4	0.428
Unknown	614	**56.5**	144	**67.0**	0.5	0.4	0.7	<0.001	0.6	0.4	0.9	0.015
Any systemic symptom												
No	750	**69.0**	186	**72.9**	1.0				1.0			
Yes	292	**26.9**	57	**22.4**	1.3	0.9	1.7	0.150	1.6	1.0	2.4	0.034
Unknown	45	**4.1**	12	**4.7**	0.9	0.5	1.8	0.833	1.5	0.6	3.3	0.360
Any other STI^												
No	664	**61.1**	143	**56.1**	1.0				1.0			
Yes	315	**29.0**	72	**28.2**	0.9	0.7	1.3	0.718	1.2	0.8	1.8	0.453
Unknown	108	**9.9**	40	**15.7**	0.6	0.4	0.9	0.010	0.5	0.3	0.9	0.014
Hepatitis C (PCR)												
No	410	**37.7**	57	**22.4**	1.0				Not included			
Yes	138	**12.7**	3	**1.2**	5.9	2.0	17.5	0.001				
Unknown	539	**49.6**	195	**76.5**	0.4	0.3	0.5	<0.001				

*Multivariable model adjusted also for form delay.

^None against gonorrhoea, syphilis, NSU, warts, herpes, hepatitis B and some less common STIs (described in data cleaning in Appendix for chapter 2, Table S 1).

2.7.1.1 Risk behaviour

In behavioural variables, presented in Table 5, the univariate analysis revealed a consistent positive association between unprotected sexual practices and being HIV-positive. The strongest association was seen in reporting both receptive and insertive unprotected fisting ("reported both, unprotected") with an OR 3.1 (95 % CI 1.4, 6.9) followed by reporting unprotected receptive anal intercourse (OR 2. 9, 95% CI 1.7, 4.8). Similarly reporting both insertive and receptive oral sex had a positive association ("reported both, unprotected" OR 1.9, 95% CI 1.2, 3.0). Also older age had a positive association with being HIV-positive. Of the sexual practices reporting protected or protection unknown IAI, unknown IAI and any vaginal sex had a negative association with being HIV-positive: OR 0.6 (95% CI 0.4, 1.0), OR 0.9 (95% CI 0.6, 1.6), and OR 0.5 (95% CI 0.1, 2.6), respectively. Any vaginal sex had only six events in the dataset, and was less likely to be reported in the HIV-positive group, making it a poor estimate in statistical analysis. Reporting one's sexuality as bisexual also has a negative association with being HIV-positive (OR 0.2, 95% CI 0.1, 0.5) however there were only 22 episodes of reported bisexual status in the dataset.

In the behavioural multivariable model I included all behavioural variables apart from ethnicity, sharing sex toys and vaginal intercourse. Ethnicity was not included as it did not reveal potential differences between subpopulations in the descriptive analysis (p-values >0.2). Sharing sex toys was not included due to previously described correlation with fisting (in section 2.3.2). Fisting has been previously described as a potential risk factor for LGV (Ward, Macdonald, Ronn, *et al.*, 2011), and can be considered as a more influential measurement of risk behaviour than sharing sex toys; furthermore the association between unprotected fisting and being HIV-positive was larger than for sharing sex toys in univariate analysis (OR 3.1, and OR 1.2, respectively). Vaginal intercourse was not included due to few events in the dataset.

Of the behavioural variables, the largest significant association for being in the HIV-positive group was reporting unprotected RAI (aOR 3.05, 95% CI 1.53, 6.11); noteworthy is also that the HIV-positive group was almost three times more likely to have information on RAI missing. Also reporting both insertive and receptive fisting, unprotected, had a positive association with being HIV-positive though this was not significant and the wide confidence interval indicates some instability in the variable (aOR 3.69, 95% CI 0.96, 14.1), which is probably due to the few number of events in the HIV-negative/unknown group. The HIV-positive group also

had fewer men who reported themselves as bisexual, and this association remained significant in the multivariable model (aOR 0.20, 95% CI 0.06, 0.66) however it was not strongly influenced by other explanatory variables and the odds ratio remained identical to that in the univariate model.

Probable country of acquisition and meeting new partners did not reveal differences in the univariate nor multivariable models, and of the sexual practices analysed, insertive anal intercourse did not have a strong association with being HIV-positive when compared to the HIV-negative/unknown group. The multivariable model had 878 observations in 842 groups, and had 31 explanatory variable levels (including categories) resulting in 27.16 groups per variable level (924/26)

Table 5. Association between behavioural factors and being HIV-diagnosed in MSM with LGV: results of univariate and multivariable analysis

	HIV+ (n=1087)		HIV-/unknown (n=255)		Univariate logistic regression (GEE)				Multivariate logistic regression* (GEE)			
	n	%	n	%	OR	1.0	CI	P-value	OR	1.0	CI	P-value
Presentation year												
Before 2010	731	**67.7**	196	**77.2**	1.0				1.0			
During 2010	349	**32.3**	58	**22.8**	1.6	1.2	2.2	0.003	1.5	1.0	2.3	0.034
Age (years, continuous)												
mean (sd)	38.60 (8.08)		37.00 (9.92)		1.0	1.0	1.0	0.023	1.0	1.0	1.0	0.072
Ethnicity												
White	953	**87.7**	229	**89.8**	1.0				Not included			
Black	51	**4.7**	9	**3.5**	1.3	0.7	2.8	0.426				
Asian	30	**2.8**	7	**2.8**	1.0	0.4	2.3	0.969				
Other	36	**3.3**	7	**2.8**	1.2	0.5	2.7	0.650				
Unknown	17	**1.6**	3	**1.2**	1.3	0.4	4.4	0.625				
Sexuality												
Homosexual	1,076	**99.0**	243	**95.3**	1.0				1.0			
Bisexual	11	**1.0**	12	**4.7**	0.2	0.1	0.5	<0.001	0.2	0.1	0.7	0.008
Acquisition country												
UK	828	**76.2**	198	**77.7**	1.0				1.0			
Abroad	81	**7.5**	16	**6.3**	1.2	0.7	2.1	0.476	1.1	0.5	2.3	0.812
Either	47	**4.3**	13	**5.1**	0.9	0.5	1.6	0.651	0.7	0.3	1.7	0.497
Unknown	131	**12.1**	28	**11.0**	1.1	0.7	1.7	0.604	0.8	0.4	1.7	0.634
Met partners in sex-on-presmises venue^												
None reported	277	**25.5**	63	**24.7**	1.0				1.0			
Met sex partners in these locations	263	**24.2**	50	**19.6**	1.2	0.8	1.8	0.371	0.8	0.4	1.3	0.302
Unknown	547	**50.3**	142	**55.7**	0.9	0.6	1.2	0.449	0.7	0.5	1.1	0.182
Number of contacts (continuous)												
mean (sd)	8.27	(16.66)	6.56	(17.18)	1.0	1.0	1.0	0.354	1.0	1.0	1.0	0.134
median (range)	3	(0-201)	3	(0-213)								
Receptive anal intercourse												
None reported	51	**4.7**	24	**9.4**	1.0				1.0			
Reported protected/or prot. unk.	161	**14.8**	79	**31.0**	0.9	0.5	1.6	0.840	1.4	0.6	3.0	0.440
Unprotected	791	**72.8**	127	**49.8**	2.9	1.7	4.8	<0.001	3.1	1.5	6.1	0.002
Unknown	84	**7.7**	25	**9.8**	1.6	0.8	3.0	0.180	2.8	1.0	8.3	0.057
Insertive anal intercourse												
None reported	100	**9.2**	25	**9.8**	1.0				1.0			
Reported protected/or prot. unk.	154	**14.2**	65	**25.5**	0.6	0.4	1.0	0.056	0.8	0.4	1.6	0.556
Unprotected	615	**56.6**	106	**41.6**	1.5	0.9	2.3	0.117	1.1	0.6	2.1	0.803
Unknown	218	**20.1**	59	**23.1**	0.9	0.6	1.6	0.811	1.1	0.5	2.6	0.736
Any oral sex												
None reported	72	**6.6**	28	**11.0**	1.0				1.0			
Reported some	23	**2.1**	6	**2.4**	1.5	0.6	3.9	0.439	1.8	0.5	6.3	0.382
Reported one unprotected	30	**2.8**	10	**3.9**	1.2	0.5	2.7	0.728	1.0	0.3	3.0	0.976
Reported both unrprotected	812	**74.7**	165	**64.7**	1.9	1.2	3.0	0.008	1.6	0.7	3.5	0.289
Some or all unknown	150	**13.8**	46	**18.0**	1.3	0.7	2.2	0.414	0.8	0.3	2.1	0.689
Any fisting												
No fisting reported	419	**38.6**	114	**44.7**	1.0				1.0			
Some fisting reported	52	**4.8**	9	**3.5**	1.6	0.8	3.2	0.224	2.1	0.7	6.3	0.170
Both reported, unprotected	71	**6.5**	6	**2.4**	3.1	1.4	6.9	0.007	3.7	1.0	14.1	0.056
Some unknown	545	**50.1**	126	**49.4**	1.2	0.9	1.6	0.245	1.1	0.7	1.7	0.655
Sharing sex toys												
No	412	**37.9**	108	**42.4**	1.0				Not included			
Any (prot or unpr)	77	**7.1**	17	**6.7**	1.2	0.7	2.0	0.581				
Unknown	598	**55.0**	130	**51.0**	1.2	0.9	1.6	0.193				
Vaginal intercourse												
No	798	**73.4**	201	**78.8**	1.0				Not included			
Yes (unprot or not)	4	**0.4**	2	**0.8**	0.5	0.1	2.6	0.415				
Unknown	285	**26.2**	52	**20.4**	1.4	1.0	1.9	0.051				

*Multivariable model adjusted also for form delay.

^No new partners in these locations against backroom, sauna, cruising ground and sex party.

2.7.1.2 Multivariable model combining clinical and behavioural factors

The purpose of the concluding multivariable model is to explore the potential joint effects of clinical and behavioural factors by including the most relevant factors together. In the final model all clinical variables, which were included in the first multivariable model, were also included in the final model. Presentation date to the clinic had no association in the multivariable clinical model (p-value 0.989), but it had a positive association in the multivariable behavioural model (OR 1.5, 95% CI 1.0, 2.3), and was therefore included in the final model. All other variables in the clinical multivariable model were considered of potential interest, and were included in the final model (they also all had a category with <0.2 p-value).

Of the variables which were in the first behavioural multivariable analysis (in Table 5), all variables except probable country of acquisition and sex-on-premises venue for meeting partners were included; these two variables did not reveal a strong association with the dependent variable when defined by the <0.2 p-value cut-off point[17] and neither were considered a priori to be important explanatory factors. They are more distal measures of LGV acquisition risk, and for sex-on-premises venue over half of the observations were in the unknown category, whilst acquisition country revealed no great differences in univariate or multivariable level. In contrast to this, although insertive anal intercourse and oral sex both had their p-values above 0.2 for all measured categories, I considered these variables as proximate measures of one's sexual risk behaviour and included them in the final model.

In the final model, presented in Table 6, of the clinical variables, longer duration of symptoms (more than a week aOR 0.5, 95% CI 0.3, 0.8) remained negatively associated with HIV-positive status. Referral as a reason for attending the clinic had a negative association with being HIV-positive (aOR 0.4, 95% CI 0.1, 1.1) and reporting any systemic symptoms had a positive association (aOR 1.6, 95% CI 1.0, 2.6), but neither is significant at 0.05 level. Having an unrecorded value for other concurrent STIs remained negatively associated with HIV-positivity (aOR 0.5, 95% CI 0.3, 1.0). In univariate analysis we saw a negative association with genital infection (site) and HIV positivity (OR 0.3, 95% CI 0.1, 0.9) while in this model the association seems to completely

[17] However meeting venue had an unknown category with p-value of 0.182, this was not included as reporting the event itself was not strongly associated with HIV status (aOR 0.75, 95% CI 0.75, 1.29).

disappear (aOR 1.2, 95% CI 0.3, 5.6); however the number of events in this category is very low (10 and 5 in HIV-positive and unknown group, respectively) and makes inferences based on statistical analysis inappropriate.

Unprotected RAI had the strongest significant association with being HIV-positive in the multivariable model (aOR 2.7, 95% CI 1.3, 5.8). Reporting unprotected fisting had a strong but non-significant association (aOR 3.5, 95% CI 0.8, 15.5). Reporting sexuality as bisexual had a negative association with being HIV-positive but the magnitude remained unchanged from the univariate model (aOR 0.2, 95% CI 0.1, 0.6). Age had a positive association with HIV-positivity but this did not reach statistical significance (OR 1.0, 95% CI 1.0, 1.1). The final model had the most variables in it with 47 explanatory levels for 878 observations in 842 groups resulting in 17.9 (842/47) groups per variables level.

2.7.1.3 Comparison of the results to standard logistic regression

Given the repeated infections in the dataset, I selected a conservative approach and used GEE instead of standard logistic regression. However the overall number of repeated episodes is relatively small and it was not clear if using a more complicated statistical model made a difference to the overall results. Therefore I performed the univariate analysis using standard logistic regression and compared the results to univariate GEE (see the appendix and Table S 6 for clinical variables and Table S 7 for behavioural variables).

Given a large sample size, and data which is not missing completely at random, GEE and standard logistic regression should produce very similar coefficient estimates (Hu et al. 1998, who refer to S. L. Zeger 1988). However, the standard errors will be different producing different confidence intervals. Hu et al. (1998) demonstrate with their data-analysis that this bias is different whether the independent variables are time-dependent; standard logistic regression overestimated standard errors for time-varying variables, and underestimated for time-invariant variables. In this dataset the result between GEE analysis and standard logistic regression are very similar. The biggest difference in effect size and the width of the confidence interval seem to be in the hepatitis C PCR results (OR 5.9, 95% CI 2.0, 17.5 in GEE compared to OR 6.4, 95% CI 2.0, 20.8 in standard logistic regression). Other risk behaviour variables, namely unprotected RAI, oral sex and fisting demonstrate small differences in the effect estimate and confidence intervals.

Table 6 Final model combining behavioural and clinical correlates

	Multivariate logistic regression* (GEE)			
	OR	1.0	CI	P-value
Presentation year				
Before 2010	1.0			
During 2010	1.1	0.7	1.8	0.736
Seen in a clinic in London				
No	1.0			
Yes	1.1	0.7	1.7	0.648
Duration of symptoms				
Week or less	1.0			
More than a week	0.5	0.3	0.8	0.002
Unknown	0.7	0.4	1.4	0.369
Referral				
No	1.0			
Yes	0.4	0.1	1.1	0.073
Unknown	0.4	0.2	1.2	0.107
Location of symptoms reported				
None	1.0			
Only Genital	0.4	0.1	1.2	0.118
Only Rectal	1.4	0.6	3.4	0.441
Both	0.8	0.3	2.0	0.567
Unknown	0.9	0.3	2.8	0.879
Site of infection				
Rectal	1.0			
Genital	1.2	0.2	5.5	0.847
Both or other (throat, n=1)	0.5	0.1	2.0	0.303
Unknown	0.6	0.4	1.0	0.053
Any systemic symptom				
No	1.0			
Yes	1.6	1.0	2.6	0.057
Unknown	2.0	0.8	5.1	0.157
Any other STI				
No	1.0			
Yes	1.1	0.7	1.7	0.601
Unknown	0.5	0.3	1.0	0.037
Age	1.0	1.0	1.0	0.099
Sexuality				
Homosexual	1.0			
Bisexual	0.2	0.1	0.6	0.006
Number of contacts	1.0	1.0	1.0	0.370
Receptive anal intercourse				
None reported	1.0			
Reported protected/or prot. unk.	1.4	0.6	3.3	0.428
Unprotected	2.7	1.3	5.8	0.010
Unknown	2.5	0.8	7.6	0.109
Insertive anal intercourse				
None reported	1.0			
Reported protected/or prot. unk.	0.8	0.4	1.6	0.473
Unprotected	1.2	0.6	2.2	0.668
Unknown	1.2	0.5	2.7	0.641
Any oral sex				
None reported	1.0			
Reported some	1.9	0.5	7.0	0.336
Reported one unprotected	1.2	0.4	4.0	0.730
Reported both unrprotected	1.9	0.9	4.2	0.108
Some or all unknown	1.0	0.4	2.5	0.992
Any fisting				
No fisting reported	1.0			
Some fisting reported	1.9	0.6	6.3	0.297
Both reported, unprotected	3.5	0.8	15.5	0.093
Some unknown	0.9	0.6	1.4	0.777

*Multivariable model adjusted also for form delay

2.8 Discussion

In this chapter I have described the structure of the diagnostic services for LGV and the data cleaning procedure for the LGV Enhanced Surveillance dataset. This provided the clinical and methodological context for subsequent analyses which examined key differences between HIV-positive and HIV-negative/unknown subpopulations with LGV, and explored reasons for the high prevalence of HIV among LGV cases.

Two datasets were available: laboratory description of samples that were positive for *C. trachomatis* and were tested for LGV, and a dataset of confirmed LGV episodes reported to LGV Enhanced Surveillance data. In the United Kingdom, testing of LGV is focused on patients with LGV-type symptoms who test positive for rectal chlamydia, with 91.1% of male samples tested for LGV being rectal swabs (and 93.3% of LGV confirmed samples being from rectal swabs). Until the end of 2010, LGV surveillance was reliant on two sources of data: laboratory testing and enhanced surveillance. Laboratory and clinical surveillance data are based on confirmed LGV diagnoses. However, the selection of patients for LGV testing is based on clinical suspicion i.e. potential asymptomatic cases would be missed using this algorithm.

The surveillance systems and previous observational studies have detected a persistent association between LGV and HIV status (Savage, van de Laar, Gallay, *et al.*, 2009; Rönn & Ward, 2011). In the data analysis presented here I used logistic regression with generalised estimating equations (GEE) to explore differences between HIV-positive and HIV-negative/unknown LGV patients.

The first objective was to look at differences in clinical presentation. Genital symptoms and genital site of infection were more common among HIV-negative/unknown LGV patients in univariate analysis but the association was no longer significant in any of the multivariable models. The scarcity of genital infection in the dataset makes drawing inferences from this challenging, and as a group HIV-negative/unknown men are not considerably different from the HIV-positive in their clinical presentation for the variables that were measured

for in the dataset[18]. Given the duration of symptoms was longer for HIV-negative/unknown patients than for HIV-positive patients, a possible explanation for the clinical differences between the comparison groups could be differential contact with health care services and a diagnostic bias in favour of the HIV positive men. This is indirectly supported by the HIV-negative/unknown men being more likely to have information missing on site of infection (aOR 0.63, 95% CI 0.39, 1.01) and concurrent STIs (aOR 0.53, 95% CI 0.29, 0.96) compared to HIV-positive men, suggesting systematic differences in how information is gathered/stored for these two groups. Alternatively, shorter duration of symptoms among HIV-positive men could be due to severity of symptoms (if HIV-positive men develop more severe symptoms and thus seek medical care more quickly).

The second objective was to look at the behavioural factors. HIV-positive men were over three times more likely to report unprotected RAI and five times less likely to report being bisexual. Reporting unprotected fisting and older age also had a positive association with being HIV-positive although this was not significant at the 0.05 significance level. The third objective was to look at the joint effects of the clinical and behavioural factors. In this multivariable model HIV-negative/unknown men were twice as likely to have symptoms for more than a week and five times more likely to be bisexual whilst HIV-positive men were 2.7 times more likely to report unprotected RAI.

In a previous case-finding exercise performed in the UK, LGV positivity in MSM attending GUM clinics was found to be 0.90% (0.69% to 1.16%) for rectal samples and 0.04% (0.01% to 0.16%) for urethral samples (Ward, Alexander, Carder, *et al.*, 2009); in chlamydia-positive samples tested for LGV at STBRL the positivity in rectal samples was 15.5% compared to 7.1% and 2.9% in urethral and urine samples. Taken together, these findings suggest that LGV may be more often acquired in the rectum than in the urethra. However, these test

[18] If HIV-positive men have different clinical pathways to HIV-negative men the HIV-negative men might be expected to have an atypical presentation compared to HIV-positive men leading the clinician to test them for LGV. I investigated the 255 HIV-negative/unknown men in more detail. Of the 255 who were HIV-unknown, 218 (85.5%) attended the clinic due to symptoms, contact tracing was mentioned for 20 (7.84%), 16 (6.27%) attended via referral and 20 (7.84%) had a routine STI screen (several options may have been selected); reasons for attending were similar for HIV-positive LGV patients. Only 5 (1.96%) of the HIV-negative/unknown patients had a genital infection and further 3 (1.18%) had both genital and rectal infections with the rest with either rectal infection (66, 25.9%) or unknown site of infection (181, 70%).

results could be biased by missing urethral samples as a result of asymptomatic infection or misdiagnosis (and treatment) as non-LGV chlamydia which may be more likely in genital versus rectal infections. (The same could apply to asymptomatic rectal infections). In the current epidemic of LGV, clinical presentation has predominantly manifested as rectal symptoms (Stamm, 2008b; White, 2009) and therefore the majority of surveillance systems in Europe have focused on detecting rectal LGV (Rönn & Ward, 2011).

Potential differences in the natural history of LGV due to underlying HIV-infection cannot be excluded based on this finding. Interestingly, a similar difference in infection site (as was seen at univariate, but not at multivariable-level) was noted in Amsterdam where half (5/10) of the LGV patients with inguinal infections were HIV-positive compared to 82.7% (340/411) of the LGV patients with rectal infections (supplement material for De Vrieze et al. 2013). Although the study suffers from a small sample size, similar to this study, the STI clinic in the Amsterdam study tests all rectal, ulcer and bubo samples positive for *C. trachomatis* for LGV regardless of clinical presentation, making the result more generalizable in that sense; however the clinic does not screen urethral chlamydia for LGV.

An alternative explanation might be that different sexual practices lead to different sites of infection: HIV-positive LGV patients report more unprotected RAI than HIV-negative/unknown men (73% compared to 50%). Whilst HIV-positive men also reported more unprotected IAI (57% versus 42%), unprotected RAI remained significant in the final multivariable model (aOR 2.7, 95% CI 1.3, 5.8) whilst unprotected IAI did not. This could indicate either general preference for receptive role that predisposes some HIV-positive men for increased risk for STI acquisition, or it could be a sign of strategic positioning whereby HIV-positive men take the receptive role and HIV-negative men the insertive role to reduce the risk of HIV transmission (Van de Ven, Kippax, Crawford, *et al.*, 2002). However, if this is the case, we would expect to see more genital infection in HIV-negative men (unless genital infection is more likely to be asymptomatic).

HIV-positive MSM, on average, report more high-risk behaviour than HIV-negative MSM in the UK (Dodds, Johnson, Parry, *et al.*, 2007). Enhanced surveillance data does not inform us of partnership or sexual network determinants. The overlap with hepatitis C infection and the information on the risk profile and hepatitis C

networks (van de Laar, Pybus, Bruisten, *et al.*, 2009) would indirectly support the idea that LGV could be transmitted in dense networks of HIV-positive high-risk men. Furthermore in our study there were few men who identified as bisexual, but they were more likely to be HIV-negative (an association which remained in the final multivariable model aOR 0.2; 95% CI 0.1, 0.6) which we may interpret as an indirect evidence of many HIV-negative LGV patients being more on the margins of the predominantly homosexual network where LGV is transmitted.

Therefore it seems that HIV-negative men who acquire LGV are a missed opportunity for primary HIV prevention, such as counselling offered at STI clinics (which at least some central London clinics offer). The HIV-negative LGV patients could also benefit from novel methods such as Pre-exposure Prophylaxis (PrEP) (Grant, Lama, Anderson, *et al.*, 2010) and they might belong to the type of high-risk group required to make PrEP cost-effective (Gomez, Borquez, Caceres, *et al.*, 2012).

2.8.1 Strengths and limitations

Despite the challenges posed by re-emergence of LGV on the health care system, the LGV diagnostics at STBRL and the LGV Enhanced Surveillance system demonstrate that surveillance can be rapidly established and produce interesting data and hypotheses. This type of analysis on the Enhanced Surveillance data offers a retrospective view on the re-emergence of LGV. The laboratory surveillance of LGV can detect changes in numer of diagnoses, but the delays incurred in LGV Enhanced Surveillance, limit its suitability for real-time outbreak analysis.

This analysis was limited in respect to looking at HIV status as the dependent variable: no confirmatory variable exist for HIV-negatives, and hence there is potential for HIV-positives to be misclassified in the HIV negative/unknown group. This is more likely to dilute any associations between explanatory variables and HIV status than strengthen them. Furthermore, unless the patient refuses, HIV testing would be performed at a GUM clinic visit as part of the STI screen (with HIV testing high in this group). Therefore the potential misclassification is more likely to be due to poor form completion of HIV-related questions (as evidenced by the 40 episodes that had discrepant information on the HIV status) rather than HIV remaining undiagnosed at a clinic visit.

LGV enhanced surveillance operated retrospectively, and therefore the forms relied on information collected from clinician's notes or from the patient during a follow-up visit. I tried to take this into consideration by including the form delay variable in all the multivariable analyses; however the retrospective data collection may still cause recall bias, especially for behavioural variables. Also, we were only able to control for repeat infections that were linked via laboratory and enhanced surveillance. The GEE method considers clustering to be a nuisance of no intrinsic interest and it estimates parameters whilst correcting for the clustering (Kirkwood & Sterne, 2008). In order to analyse the potential impact of repeat infections on the LGV epidemic, I will look at the individuals with more than one infection in more detail in the next chapter.

Another important consideration is the relatively small number of observations in the comparison group (19% of the dataset in the analysis). This limits the number of variables feasibly included in multivariable model as discussed by Kirkwood and Sterne (2008) and Hosmer and Lemeshow (2000). The overall number of groups per explanatory level was sufficient (for the largest model 17.91 groups per variable level), however if we consider the number of groups in the HIV –negative/unknown group as the limiting factor (n=254), the final model has only 5.40 HIV-unknown groups per explanatory variable level (254/47). One option that was discussed during the analysis was to reduce the number of categories in the variables further. I decided against this as this also entails loss of information, which may increase the amount of residual confounding due to reduced accuracy in describing the data. I also considered the complexities of the dataset of interest, and therefore wanted to maintain variables with more levels giving transparency to interpretation.

2.9 Conclusions

In light of the hypotheses, longer duration of symptoms would indicate issues in the diagnosis of LGV with HIV-negative men taking longer to be diagnosed. On the other hand the persistent positive association between unprotected sexual practices and HIV-positivity supports the concept of differential risk behaviour, which could be due to differences in risk taking or more specifically due to serosorting and other seroadaptive behaviours. Unprotected RAI remains strongly associated with the HIV-positive group when controlled for other explanatory variables, and this being an important acquisition risk factor is plausible given the majority

of confirmed LGV is rectal. It could also be that different sexual practices (or different natural history depending on underlying HIV infection) led to the site of infection being different for HIV-positive and HIV-negative/unknown MSM. However the evidence for this is not very strong. We cannot exclude the possibility of increased biological susceptibility for LGV in HIV-positive men based on this study. But if the underlying susceptibility due to HIV-infection was a driver of the differential distribution of LGV, we might expect there to be fewer differences between HIV-positive and HIV-negative/unknown LGV cases across several risk behaviours measured in the data.

However it does seem the HIV-positive LGV patients, in general, report higher level of risky sex, and in particular unprotected receptive anal intercourse, which is a plausible route of acquisition. Also reporting fisting had a non-significant positive association, and it has been suggested as a possible risk factor (de Vries, van der Bij, Fennema, *et al.*, 2008). Based on few people reporting the practice it is possibly a marker of risk-behaviour or network membership in general.

2.10 Recommendations

The coverage of the Enhanced Surveillance system has been good, which likely reflects the centralised nature of LGV testing in STBRL. The surveillance described the key population at risk for LGV, and has demonstrated a stable epidemiological profile of LGV cases across data collection, and lymphogranuloma venereum seems to remain confined to a small subpopulation of HIV-positive MSM.

However the surveillance may be systematically missing asymptomatic cases – as demonstrated by a recent case-finding exercise - as well as there being a delay in recognising cases in HIV-negative men as indicated by the data-analysis in this chapter. This is a result of the protocol used for testing together with clinical requests for tests. Given the current evidence, the algorithm of LGV testing should be modified in response to this to be able to capture asymptomatic and HIV-negative LGV cases. This can be done by analysing the case-finding data, and other available epidemiological data to identify risk factors for LGV, and target LGV testing accordingly. Alternatively all rectal samples positive for chlamydia in MSM could be tested for LGV but this would incur a significant additional burden on the diagnosing laboratory, STBRL.

As the Enhanced Surveillance ended in 2010, the LGV surveillance currently relies on laboratory surveillance by the STBRL, and we are lacking further information on the epidemiology of LGV. Given the current low level of LGV, enhanced surveillance may no longer be practical, but LGV surveillance might benefit from supplemented periodic case-finding and collection of detailed information on patients to monitor the epidemiology of the disease. This can also aid in monitoring the potential spread of LGV outside the core population. It would also allow estimation of LGV positivity among MSM. LGV case-finding data would have been an alternative source of data for the analysis in this chapter, had the data existed at the time.

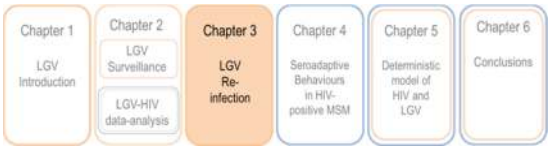
The surveillance did not measure recreational drug use, and problems related to that, and unforeseen aspects of risk behaviours can be discovered by qualitative research, a method that has received little attention in LGV research. Qualitative research has proven particularly useful for shigella outbreaks in MSM, and similar approaches could be applied to LGV surveillance; both being rare infections among MSM, understanding of the specific risk behaviours becomes important for appropriate control measures and prevention messages.

Finally the accumulating evidence should be used to design targeted prevention messages that are suitable for the key population affected. As an example LGV control measures should include outreach and direct contact to key venues and social networks, such as the internet.

Chapter 3

Core within a core?

Analysis of LGV patients with a known repeat infection



3.1 Summary

In the previous chapter I investigated the differences in HIV-positive and HIV-negative/unknown LGV patients. I used generalised estimating equations (GEE) to control for the impact of re-infections. Determining reasons for what makes an LGV patient more likely to acquire LGV in the future is of importance in designing effective interventions. In this chapter I aim to identify characteristics that might explain re-infection among the men in the dataset by comparing baseline characteristics of men with repeat LGV episodes to men with a single reported episode. This is done through logistic regression or Fisher's exact test where the number of events is few.

In this chapter I will show that those who had a reported LGV re-infection were more likely to be HIV-positive, visit a clinic in London and had hepatitis C and concurrent gonorrhoea on their first recorded LGV episode. Repeaters also reported higher levels of unprotected sex but this was not statistically significant. Due to low number of individuals with repeat infection, it was not possible to create a predictive model for probability of future infection. Furthermore behavioural variables were not markedly different between repeaters and non-repeaters.

LGV repeaters display characteristics traditionally attributed to core groups but behaviour alone did not explain the later occurrence of re-infection. LGV repeaters have a high prevalence of STI co-infection which is of clinical and epidemiological relevance, and may suggest their position in the sexual network is contributing to their heightened risk of STI acquisition. Given the stochastic nature of LGV occurrence underlying LGV prevalence may be a more determining factor in a population where the overall patterns of risk are elevated. Further research is needed of sexual network structures.

3.2 Introduction

Core groups have become a central feature of conceptualising the determinants of STIs with host heterogeneities thought to be the main driver of infection persistence; infection saturation occurs in subpopulations with higher transmission-related risk behaviour, resulting in lower equilibrium prevalence but higher reproductive number than if the whole population had the same average behaviour patterns (Keeling & Rohani, 2008; Anderson & May, 1991). Core groups can be defined as those individuals who transmit an infection to more than one new host (reflecting their function in maintaining reproductive number above 1) (Ghani & Garnett, 2000) or as subpopulations who are small in size but whose STI rates surpass that of the general population (reflecting acquisition and burden of STIs in these groups) (Fenton, Mercer, Johnson, *et al.*, 2005).

To identify core groups based on transmission is challenging as the direction of transmission is difficult to establish. Thus acquisition of infection is often used as a proxy measure. Partner tracing and sexual network structures can be used to identify individuals who are central to sexual networks, as was done with syphilis outbreak in North Carolina (Doherty, Adimora, Muth, *et al.*, 2011), and sexual networks can also identify individuals who connect different components together and who, if infected, have greater potential for onwards transmission (De, Singh, Wong, *et al.*, 2004). Molecular typing has also been used for HIV (Lee, Tam, Tan, *et al.*, 2009), hepatitis C (van de Laar, Pybus, Bruisten, *et al.*, 2009) and gonorrhoea (Choudhury, Risley, Ghani, *et al.*, 2006) to illustrate sexual network structures, and when this is coupled with epidemiological data on the infected this can be a powerful tool in understanding behaviour in a network context.

The prevention argument for the core group definition is to be able to target interventions to individuals who are at high-risk for STI transmission, and who can be grouped by distinct characteristics (such as involvement in sex work, sexual orientation, ethnicity, geographical location, repeat STI infections) (Thomas & Tucker, 1996). The approach has been criticised for concentrating on the behaviour of an individual and ignoring the social context in which risky sex takes place (Rothenberg, Potterat & Woodhouse, 1996). Core group definition can also be seen as stigmatising of an entire group of people, and being unjust towards individuals in a given group

not taking part in high-risk behaviour (Thomas & Tucker, 1996). Alternative, less pejorative definitions, such as key groups or key populations, are also used in the literature to describe vulnerable and at risk groups in general, and this has become imperative for HIV where the infection is life-long. For instance in their terminology guidelines UNAIDS defines key populations as those who are at most risk of being exposed or to transmit HIV, and their list of key populations includes groups that were seen as traditional core groups with the addition of people who are likely to acquire but not necessarily transmit the virus (such as seronegative people in discordant partnerships). Along the same lines as the arguments on social context given by Rothenberg *et al.* the UNAIDS guidelines also object to using high-risk group as this would indicate the risk is confined within a group when in fact all groups are interconnected (UNAIDS, 2011).

Those with a repeat infection are sometimes automatically seen as members of the core group, whilst the more conservative approach would be to formulate that some members of the core group will have a repeat infection (Thomas & Tucker, 1996). However for infections where repeat infection is possible (without lifelong infection or immunity) it is likely that those who acquire infection more than once are also contributing to the onward transmission of the disease, and characteristics of those who are likely to have a repeat infection has been investigated in previous studies. In STI clinic based study in San Diego Gunn *et al.* (2000) history or current diagnosis of gonorrhoea or chlamydia were predictive of subsequent STI diagnosis. Also the more past STIs the patient reported, the higher the risk for later STI, which the researchers interpreted as indication of these people being central in sexual networks where gonorrhoea and chlamydia are transmitted. Similar findings came out from a multi-centre study from United States where women and men diagnosed with an STI were at increased risk of re-infection with the same STI but also with other STIs during a 3-month follow-up (Peterman, Tian, Metcalf, *et al.*, 2006).

Logistic regression has been used to compare baseline characteristics (from the first infection) of repeaters to those who do not experience a repeat infection, and this has been used to create a predictive model to estimate risk for future STIs in patients who had visited STI clinic in Florida (Richert, Peterman, Zaidi, *et al.*, 1993) where past STI diagnosis was the strongest predictor of new STI within a year (aOR 2.9 95% CI 2.7, 3.1) and for a repeat syphilis in MSM in San Francisco where HIV-positivity was associated with re-infection (OR

4.7; 95% CI 1.8, 12.0) (Phipps, Kent, Kohn, *et al.*, 2009). In another Californian MSM syphilis study by Cohen et al. (2011) repeat primary syphilis was associated in the multivariate analysis with HIV-positivity (aOR 1.65, 95% CI 1.14, 2.37), black (compared to white) race (aOR 1.84, 95% CI 1.12, 3.04) and having 10 or more partners (aOR 1.98, 95% CI 1.12, 3.50).

Similar analysis is done in this chapter to explore factors associated with repeat LGV infection, and factors that may help us understand their role in sustaining the epidemic.

3.3 Objectives

In this chapter I aim to describe what characterises re-infections and those who acquire them and whether we can see an indication of re-infections having a key role in sustaining the epidemic.

Two objectives were specified:

- Describe the repeaters and assess how their episodes are distributed through data collection period.
- Compare the characteristics of those with a documented repeat infection to the episodes of non-repeaters.

3.4 Methods

3.4.1 Methods to analyse repeat infections in surveillance data

In addition to applying standard logistic regression to a surveillance dataset, they have also been viewed as retrospective cohorts, and analysed as survival analysis, as done with repeat chlamydial infection in women in Australia (Batteiger, Tu, Ofner, *et al.*, 2010), chlamydia in Denver, US (Rietmeijer, Van Bemmelen, Judson, *et al.*, 2002) and repeat gonorrhoea infection in Sheffield (Hughes, Nichols, Peters, *et al.*, 2012). However as survival analysis compares time to an event between groups we need to assume that individuals have been "followed" after primary event, which in turn requires a dataset where the coverage of surveillance system is high with strong likelihood of cases and repeat infections being correctly detected. In the publications above, testing for the infection has been part of routine screening so that the coverage can be assumed to be good (Batteiger, Tu, Ofner, *et al.*, 2010; Hughes, Nichols, Peters, *et al.*, 2012) or they also had information on

negative tests which gives a better estimate of a patient's infection status during a given follow-up (Rietmeijer, Van Bemmelen, Judson, *et al.*, 2002). A systematic review of chlamydia and gonorrhoea re-infection in men by Fung et al. (2007) compared prospective and retrospective cohorts, and found that in prospective cohort studies (active follow-up) the follow-up time varied from 10-24 weeks (median 4 months) whilst in retrospective cohorts (passive follow-up) the follow-up time was much longer with a maximum of 4.8 years reported. This will affect the re-infection proportion and incidence estimates in the sample. Active follow-up can also measure loss to follow-up in their study population which passive follow-up is unable to do; however at least one prospective cohort study did not find differences in factors associated with chlamydia infection between those who were followed-up and those lost to follow-up (Lamontagne, Baster, Emmett, *et al.*, 2007) . Furthermore cohort studies based on clinic populations can overestimate the risk for the general population, and symptomatic individuals are more likely to attend than asymptomatic (Kent, Chaw, Kohn, *et al.*, 2004). In an analysis of chlamydia re-infection rate, the authors nicely demonstrate the sensitivity of the estimate to denominator chosen and what follow-up time was used (Torrone, Satterwhite, Scholes, *et al.*, 2013).

For LGV there is evidence that diagnosis is based on clinical suspicion, and the surveillance system is unlikely able to capture all infections. Furthermore we have data only of repeated visits where LGV was diagnosed and no information on visits where LGV was not diagnosed. Considering these factors I concluded survival analysis to be an inappropriate method for analysis of the LGV Enhanced Surveillance dataset, and chose the somewhat simpler logistic regression approach instead as the primary means of analysis. The benefit of logistic regression is that we are able to look at the data 'as it is' without having to make many further assumptions of the sample the data arises from (the limitation of non-differential classification bias of repeat infection status which will be discussed later). The weakness of logistic regression in this context is that we lose the aspect of time to re-infection and cannot estimate rate of infection.

3.4.2 Identifying repeaters

To identify repeaters in the LGV Enhanced Surveillance dataset patient records were linked where possible. Mostly this was done within clinic but in few instances (five individuals) between clinics where those clinics were under the same NHS trust. In short, LGV ES data and the datasheet for linkage was used as the main source for determining episode's status as part of repeat infection series or not (also briefly described in

chapter 2 in section 2.3.3). The full explanation of repeat episodes and information on missing data regarding episodes that occurred in repeaters is placed in Appendix for chapter 3.

3.4.3 Statistical analysis

Univariate logistic regression was performed comparing the first episode of repeaters to non-repeaters' episodes. The parameter estimates for logistic regression are presented only for those categories where the number of events in a cell is 5 or more and in some instances this affected the reference category used. For variables where fewer than 5 events occur in any given cell Fisher's exact test is performed.

Fisher's exact test is used to calculate exact probabilities of a 2x2 table when there are few events in any of the cells. The general recommendation for chi-squared tests is that the results become unreliable when the smallest expected value is less than 5 (Kirkwood & Sterne 2008, Chapter 17) which here was simplified by using minimum of 5 events as a rule[19]. An extension of Fisher's exact test can be used for tables larger than 2x2 (with the computational disadvantage of the number of possible row and column possibilities growing). STATA calculates Fisher's exact p-value for 2x2 tables (both one- and two-sided) and for larger tables STATA calculates one-sided p-values (STATA, n.d.), which cannot be directly compared with two-sided p-value as hypergeometric sampling distribution is rarely symmetric. However, as the one-sided p-value assumes a direction of association, a simple and conservative approach is to use cut-off value of 0.025 (Agresti 2002, p. 93; Altman et al. 2000 p.157-157).

Therefore I performed univariate logistic regression where possible, but if some of the categories had less than 5 events, the parameter estimate for these was not presented, and Fisher's exact test p-value was calculated for the variable. For 2x2 tables I display two-sided p-values and for larger tables one-sided p-values.

[19] Fisher's exact test calculates the probability that the observed table or more extreme tables (with the same row and column totals) could occur by chance by using hypergeometric distribution. To obtain a one-tailed p-value tables that show differences in the same direction as that observed are included, whilst for two-tailed p-value a number of algorithms are possible to count extreme tables to both directions from that of observed (Agresti 2002, p. 91-93).

Kruskall-Wallis test was used to test differences in medians in duration between infections by year of clinic presentation (of a repeat infection). Kruskall-Wallis is a non-parametric test to estimate if the medians of the groups arise from identical distributions (which may differ by a constant amount). When two groups are compared to each other Wilcoxon rank sum test is used, which gives identical result as Kruskall-Wallis applied to two groups (Kirkwood & Sterne, 2008).

Analyses were performed in STATA SE/11.2 and figures were created with STATA SE/11.2 and Microsoft Excel.

3.5 Results

3.5.1 Occurrence of repeat infections in LGV Enhanced Surveillance

The same 1342 episodes in 1281 MSM that were included in the analysis in chapter 2 formed the basis of the dataset used here. Altogether 66 individuals who were known to be re-infected had at least one episode in the dataset[20] and there are patients who only have some of their LGV episodes in the dataset (see Appendix for chapter 3 for more information on missing data).

The occurrence of repeaters follows broadly the overall pattern of episodes in LGV Enhanced Surveillance data as presented in Figure 1. We observe a relatively steady pattern of repeaters' first episodes occurring across the surveillance period, while repeaters' second episode shows more variation. The largest number of episodes occurred in 2010 along with the largest number of repeat infections including 5/6 of third LGV episodes identified. However 2010 is also the end of the surveillance and therefore it has the longest "follow-up" time for re-infections to have a chance to occur.

To explore the impact of follow-up time on the occurrence of repeat infections I plotted the proportion of repeat infections over time in Figure 2 where the yearly number of MSM episodes in LGV Enhanced Surveillance is presented together with the proportion of episodes known to be repeat infections as red line. Looking at this it would seem like the proportion of re-infections varies over time and is the highest in 2010

[20] Repeat infections occur in the dataset in the following way: for 9 individuals there is one episode, for 53 individuals there are 2 episodes and for 4 individuals there are 3 episodes (altogether 127 episodes).

with 9.58% of all episodes being re-infections (red line in the figure). However this approach ignores the time to re-infection, and when time to re-infection is capped (orange and blue lines) the proportion of re-infections in 2010 rapidly decline.

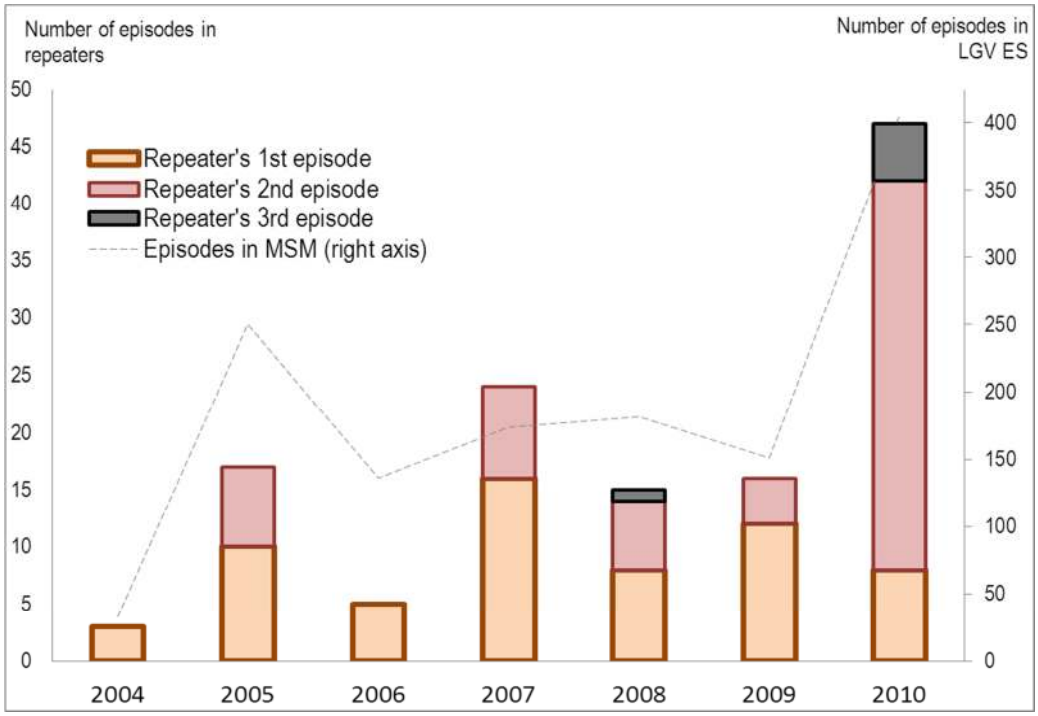


Figure 1 Number and proportional distribution of episodes that occurred in repeaters stratified by episode number and year of clinic presentation. For comparison with the overall trend the number of episodes is on the right vertical axis.

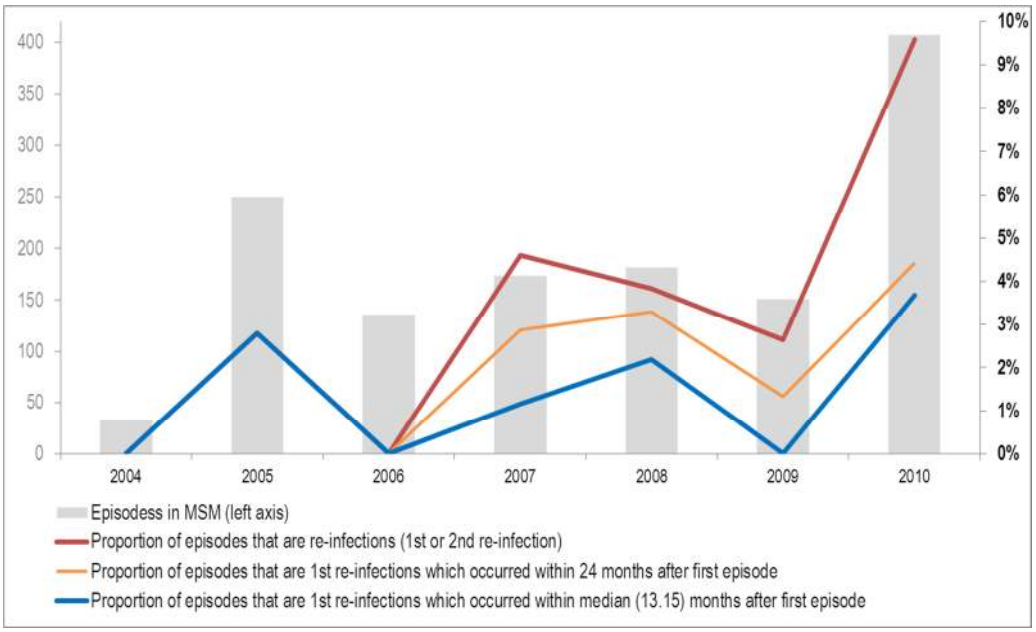


Figure 2. Yearly distribution of episodes in the LGV Enhanced Surveillance dataset. Number of episodes is on the left vertical axis and proportion of re-infections is presented on the right vertical axis.

The distribution of duration to re-infection is presented in Figure 3 where a declining distribution is seen between proportion of individuals and time to first re-infection. The median duration between the first and second episode was 13.15 months (range 3.3-51.17 months) and between second and third 32.25 months (17.53-34.40). There are very few observations to estimate time to second re-infection (3rd episode, n=4 with 2nd and 3rd episode in the dataset), but the median duration is 19 months longer (p-value 0.047 for difference using Wilcoxon rank sum).

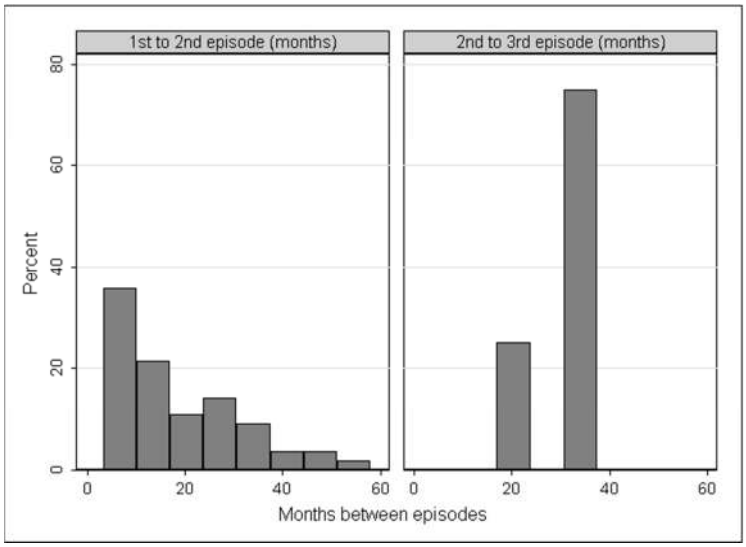


Figure 3. Distribution of time between episodes from 1st to 2nd episode (n=56) and from 2nd to 3rd episode (n=4).

More interestingly time to second LGV infection seemed to vary over the years. To test whether there is a difference in time to 1st re-infection by year of re-infection Kruskall-Wallis test was performed, and to take into account the overestimation of repeat LGV incidence when the time to re-infection is not defined, p-values are presented for various follow-up times in Table 1. When all re-infection which occurred during data collection period are included there is statistically significant variation in time to re-infection between the years (p-value 0.024). The yearly variation remains significant when the median (13.15 months) is used as the cut-off value (p-value 0.042) but is no longer significant when shorter follow-up times are used.

Table 1. Comparison of differences in time to re-infection stratified by year when the follow-up time to re-infection is varied.

Yearly variation between time to re-infection		
Second infection within	Kruskall-Wallis p-value for difference between years	Number of observations used
6 months	0.059	11
12 months	0.092	26
13.15 monts (median)	0.042	28
18 months	0.011	32
24 months	0.011	38
All included	0.024	56

3.5.2 Comparison of repeaters to non-repeaters

The next step was to assess potential predictors for future repeat infection and this was done by comparing the baseline characteristics (from 1st episode) of repeaters (n=62) to the episodes of non-repeaters (n=1,215). Four known repeaters did not have their primary episode in the database and they were not included in this analysis (see Appendix for chapter 3 for information on missing episodes). Due to low number of events in the repeaters' group, I limited the analysis to univariate level, and so robust conclusions cannot be drawn. Full list of variables analysed is presented in Appendix (Table S3) and summary of variables discussed here is presented in Table 2. Only the first episode is included in the analysis but 2nd and 3rd episodes are presented in the table and figures as well for comparison.

One of the strongest predictors for being a future repeater was HIV-positivity (only two repeaters' first episodes occurred in the HIV-negative/unknown group) with Fisher's exact p-value <0.001. Repeaters' clinic location was twice as likely to be in London as out of London (OR 2.0, 95% CI 1.1, 3.8). Of risk-factors at the time of LGV diagnosis concurrent gonorrhoea was twice as likely on repeater's first episode as on non-repeater's episode (OR 2.2, 95% CI 1.2, 3.8).

Both current hepatitis C infection (PCR) and hepatitis C antibody positivity (previous or chronic infection) were over twice as likely to occur on a repeater's first LGV episode compared to non-repeaters LGV episode. Interestingly the proportion that were hepatitis C PCR positive increased from 19.4% (12/62) on a repeater's first episode to 27.1% (16/59) and 50.0% (3/6) on repeater's second and third episode, respectively.

Table 2. Summary of variables analysed. Comparisons are between repeater's first episode and non-repeaters.

	Non-repeaters		Repeaters' 1st episode		Repeaters' 2nd episode		Repeaters' 3rd episode		Univariate logistic regression 1st episode versus non-repeaters				Fisher's exact test 2-sided/ 1-sided	
	n= 1215	%	n=62	%	n=59	%	n=6	%	OR	1.0	CI	p-value	2x2 tables/	larger tables
Age														
mean (sd)	38.2 (8.4)		38.8 (9.2)		39.4 (9.2)		44.0 (6.2)		1.0	1.0	1.0	0.618		
[range]	[18-67]		[20-62]		[21-62]		[39-56]							
HIV status														
Negative/Unknown	252	**20.7**	2	**3.2**	1	**1.7**	0	**0.0**	N/A				<0.001	
Positive	963	**79.3**	60	**96.8**	58	**98.3**	6	**100.0**						
Seen in a clinic in London														
No	394	**32.4**	12	**19.4**	10	**17.0**	0	**0.0**	1.0					
Yes	821	**67.6**	50	**80.7**	49	**83.1**	6	**100.0**	2.0	1.1	3.8	0.034		
Presentation day														
2010	847	**70.2**	54	**87.1**	25	**42.4**	1	**16.7**	1.0			0.006		
<2010	360	**29.8**	8	**12.9**	34	**57.6**	5	**83.3**	0.3	0.2	0.7			
Sexuality														
Homosexual	1,192	**98.1**	62	**100.0**	59	**100.0**	6	**100.0**	N/A				0.623	
Bisexual	23	**1.9**	0	**0.0**	0	**0.0**	0	**0.0**						
Referral														
No	1,144	**94.2**	58	**93.6**	58	**98.3**	6	**100.0**	N/A					0.023
Yes	47	**3.9**	0	**0.0**	0	**0.0**	0	**0.0**						
Unknown	24	**2.0**	4	**6.5**	1	**1.7**	0	**0.0**						
No other STIs														
No	456	**37.5**	31	**50.0**	21	**35.6**	1	**16.7**	1.0					
Yes	691	**56.9**	27	**43.6**	35	**59.3**	5	**83.3**	0.6	0.3	1.0	0.040		
Unknown	68	**5.6**	4	**6.5**	3	**5.1**	0	**0.0**	N/A					
Gonorrhoea														
No	950	**78.2**	40	**64.5**	44	**74.6**	5	**83.3**	1.0					
Yes	198	**16.3**	18	**29.0**	13	**22.0**	1	**16.7**	2.2	1.2	3.8	0.009		
Unknown	67	**5.5**	4	**6.5**	2	**3.4**	0	**0.0**	N/A					
Hepatitis C (PCR)														
No	422	**34.7**	21	**33.9**	23	**39.0**	1	**16.7**	1.0					
Yes	110	**9.1**	12	**19.4**	16	**27.1**	3	**50.0**	2.2	1.0	4.6	0.038		
Unknown	683	**56.2**	29	**46.8**	20	**33.9**	2	**33.3**	0.9	0.5	1.5	0.588		
Hepatitis C (Ab)														
No	800	**65.8**	35	**56.5**	29	**49.2**	1	**16.7**	1.0					
Yes	168	**13.8**	17	**27.4**	25	**42.4**	3	**50.0**	2.3	1.3	4.2	0.006		
Unknown	247	**20.3**	10	**16.1**	5	**8.5**	2	**33.3**	0.9	0.5	1.9	0.832		
Number of contacts														
median (range)	3 (0-213)		3 (0-85)		4 (0-100)		8.5 (3-20)		1.0	1.0	1.0	0.237		
Receptive anal intercourse														
None reported	75	**6.2**	0	**0.0**	0	**0.0**	0	**0.0**	N/A					0.029
Reported prot./ prot. unk.	225	**18.5**	9	**14.5**	6	**10.2**	0	**0.0**	0.6	0.3	1.3	0.219		
Unprotected	810	**66.7**	51	**82.3**	51	**86.4**	6	**100.0**	1.0					
Unknown	105	**8.6**	2	**3.2**	2	**3.4**	0	**0.0**	N/A					
Insertive anal intercourse														
None reported	109	**9.0**	4	**6.5**	9	**15.3**	3	**50.0**	N/A					0.380
Reported prot./ prot. unk.	204	**16.8**	9	**14.5**	6	**10.2**	0	**0.0**	0.7	0.3	1.5	0.361		
Unprotected	642	**52.8**	40	**64.5**	36	**61.0**	3	**50.0**	1.0					
Unknown	260	**21.4**	9	**14.5**	8	**13.6**	0	**0.0**	0.6	0.3	1.2	0.118		
Any fisting														
No fisting reported	485	**39.9**	26	**41.9**	21	**35.6**	1	**16.7**	N/A					0.639
Some fisting reported	54	**4.4**	3	**4.8**	4	**6.8**	0	**0.0**						
Both reported, unprotected	64	**5.3**	5	**8.1**	7	**11.9**	1	**16.7**						
Some unknown	612	**50.4**	28	**45.2**	27	**45.8**	4	**66.7**						
Any oral sex														
None reported	98	**8.1**	1	**1.6**	1	**1.7**	0	**0.0**	N/A					0.301
Reported some	26	**2.1**	1	**1.6**	2	**3.4**	0	**0.0**						
Reported one unprotected	38	**3.1**	1	**1.6**	1	**1.7**	0	**0.0**						
Reported both unrprotected	869	**71.5**	51	**82.3**	51	**86.4**	6	**100.0**						
Some or all unknown	184	**15.1**	8	**12.9**	4	**6.8**	0	**0.0**						
Sharing sex toys														
No	478	**39.3**	23	**37.1**	17	**28.8**	2	**33.3**	1.0					
Any (prot or unpr)	79	**6.5**	9	**14.5**	6	**10.2**	0	**0.0**	2.4	1.1	5.3	0.036		
Unknown/ missing	658	**54.2**	30	**48.4**	36	**61.0**	4	**66.7**	1.0	0.5	1.7	0.849		

N/A, not applicable (parameter estimate not presented due to low number of events in the cell, and category excluded from the analysis)

None of the repeaters reported not having had RAI, and 14.5% reported protected or protection unknown RAI whilst 82.3% of repeaters reported unprotected RAI. Univariate logistic regression was performed using “Some RAI” as the reference group (due to zero cell in no RAI), but unprotected RAI is not statistically different in this comparison. One-sided Fisher’s exact test was also performed for the variable and resulted in p-value above 0.025 (p-value 0.029).

To explore these trends scatter plots were created where the episodes were plotted against time (presentation date to the clinic) and stratified by episode number. Figure 4 presents episodes stratified by HIV status where the predominance of HIV-positives is clear. The figure shows how HIV-positive individuals form the majority of episodes in the LGV Enhanced Surveillance with repeat infections occurring on HIV-positives apart from one episode. Though majority of infections have occurred in London, as illustrated in Figure 5, but the distribution of repeat infections is not as skewed as for HIV.

Even though most repeaters reported unprotected RAI, so did most of the non-repeaters as can be seen in Figure 6 where few episodes in general reported no RAI prior to diagnosis with LGV. Other types of sexual behaviour were also investigated, and a trend was noted of repeaters reporting more unprotected sex in general; this is visualised in Figure 7 and displayed in Table 2. However, apart from sharing sex toys, none of the sexual behaviours achieved statistical significance in the analysis as independent variables, and non-repeaters reported high-levels of unprotected sex as well.

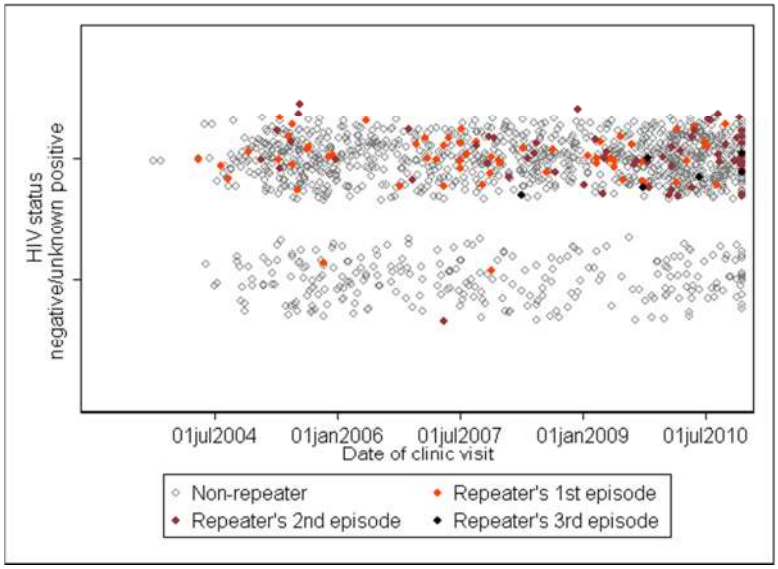


Figure 4. Distribution of episodes over time by HIV status.

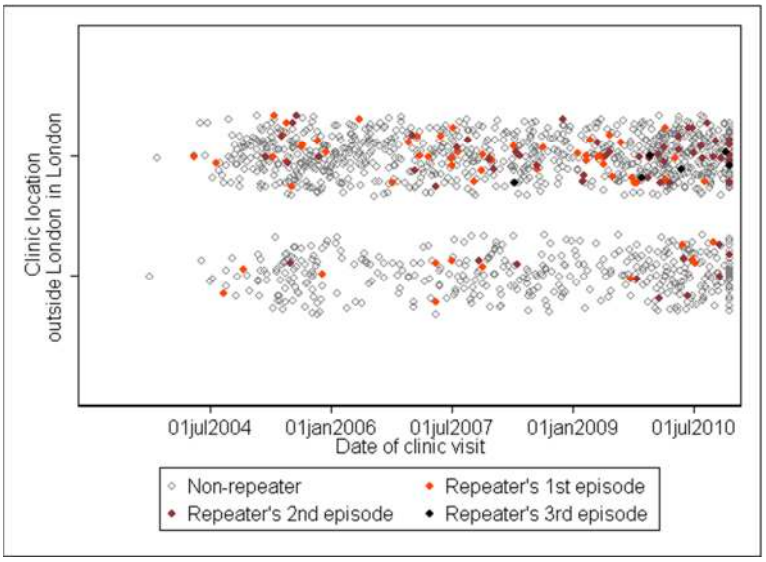


Figure 5. Distribution of episodes over time by clinic location.

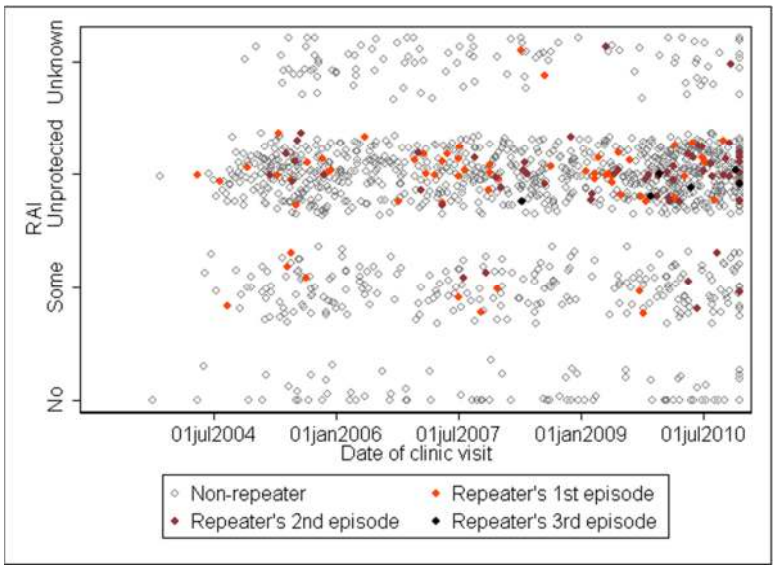


Figure 6. Distribution of episodes over time by type of receptive anal intercourse (RAI) reported

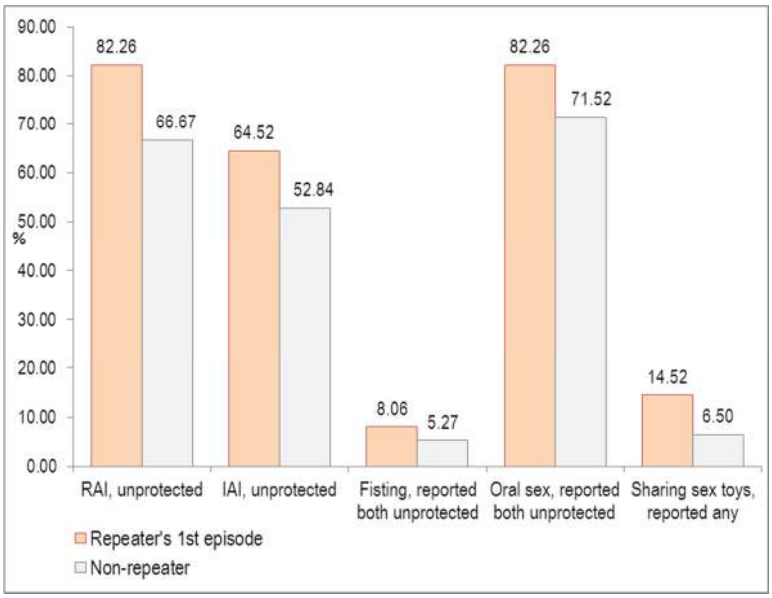


Figure 7. Proportion of LGV patients reporting unprotected sexual practices; comparing non-repeaters to repeaters' first episode.

3.6 Discussion

In this chapter I have described individuals with a known repeat infection, time to re-infection and compared repeater's first episodes against episodes without a repeat infection.

When I examined the characteristics of repeaters by comparing their first episode to non-repeaters being HIV-positive and being diagnosed in London were strongly correlated with repeat infection. Acute or chronic/previous hepatitis C infection was also associated with a repeater's first visit as was concurrent gonorrhoea infection. This would support the hypothesis of LGV repeaters being well connected to sexual networks where LGV and other STIs circulate.

The highest proportion of re-infections (9.58%, if all re-infections are considered) occurred in 2010 which also had the largest number of confirmed LGV episodes to date. Time to second re-infection seemed to be longer than time to first re-infection (p-value 0.047) and interestingly there was variation in median time to first re-infection when stratified by year, however this was strongly dependent on the timeframe in which repeat infections were considered to have occurred. Unlike in the logistic regression, where the first episodes of repeaters and non-repeaters were compared, here we are examining time to second infection which brings us back to the "passive follow-up" problem presented by Gunn *et al.* (2000). For example the longer time (and very few episodes) to second re-infection may be an indication of increasing loss-to follow up as the likelihood of patients going to a different STI clinic increases over time.

Even in the logistic regression when we are assigning individuals as repeaters or non-repeaters the assumption of time to re-infection is implicitly incorporated as the repeaters in this analysis are "ever repeated". In (Phipps, Kent, Kohn, *et al.*, 2009) syphilis re-infection was defined as second syphilis infection within a year whilst in (Cohen, Chew Ng, Katz, *et al.*, 2011) second infection had to occur within two years with an argument that time to re-infection was an indication of someone's centrality to sexual networks where syphilis is transmitted.

Cohen *et al.* (2011) estimated that 5.7% of MSM in California had a repeat syphilis infection within two years of their first infection. In the LGV Enhanced Surveillance dataset 5.2% (66/1281) of individuals had a documented re-infection in the Enhanced Surveillance if the follow-up is considered as the entire data collection period. Like with syphilis studies in MSM, HIV-positivity was associated with re-infection (Phipps, Kent, Kohn, *et al.*, 2009), However behavioural factors were not associated with LGV re-infection, but concurrent gonorrhoea and hepatitis C were. Case series of hepatitis C seroconversion in Belgium (Pelgrom, Vogelaers & Colle, 2008) and case series of LGV patients and their contacts (Gotz, van Doornum, Niesters, *et al.*, 2005) indicated an overlap between sexual networks that transmit LGV and hepatitis C. In a study by Gunn *et al.* (2000) the authors also speculated that elevated risk for any STI given a past STI diagnosis is an indication of a network position.

3.6.1 Strengths and limitations

In this analysis I had no information of loss to follow-up neither for repeaters nor for non-repeaters. Those who were captured by the enhanced surveillance system and who had the longest follow-up times (maximum of over 4 years between first and second LGV infection) can be systematically different than those who had a repeat infection within short timeframe, those who had a repeat infection which was not linked or those who had a repeat infection but LGV Enhanced Surveillance form was not filled for them. There are also LGV patients in the dataset whose repeat infection occurred after the data collection period. Unless there is an association between the exposure variables and probability of being detected as a repeater, classifying those who had a re-infection as non-repeaters would dilute the association seen. HIV-status is an exposure variable that may be associated with greater likelihood of being identified and linked as a repeater due to HIV-positive individuals having increased contact with the health care system.

As with other similar studies, here the repeat diagnosis is a proxy measure for repeat infection. HIV-positivity, time of presentation to the clinic (increased awareness over time) and whether the patient had an LGV Enhanced Surveillance form filled or not (as some repeat infections only had information for some of their episodes) all contribute to the skewed picture presented here. General trends were identified in repeaters that separated them from non-repeaters; however it would be difficult to estimate who is likely to repeat at individual-level based on this study alone. The sample size was small for repeaters which lowered the

statistical power to detect differences even if they were present. Also, it is likely that re-infections are under-reported as linking of infections occurred mainly at clinic level.

Risk for re-infection may also be determined by extrinsic factors, mainly by the prevalence of infection: as LGV is still a very rare infection even in high-risk settings; one's risk behaviour and sexual network may contribute to an extent but prevalence and stochastic events may play even bigger role, which could partially explain why the non-repeaters with high risk behaviour have not yet acquired a repeat infection with LGV..

3.7 Conclusions

If the assumptions present in this study are accurate, repeaters are part of the core group and can contribute to the persistence of LGV emergence. These findings support the hypothesis of LGV being confined to dense sexual networks of mainly HIV-positive MSM. Furthermore the concurrent STIs are of clinical importance, and are the clearest indicator of a patient's elevated risk for future LGV infections as well as risk for other STIs. Increased frequency of testing may be appropriate means of infection management in these cases. HIV-positivity, being seen in London and variation in time to re-infection between years may also be hypothesised to be proxy measures of LGV prevalence in this subpopulation.

Chapter 4

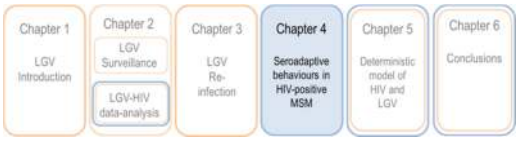
Seroadaptive behaviours in HIV-positive men who have sex with men

Proposing a conceptual framework

"I'm not going to lie…I always think, oh well, you might as well not bother; he's positive, and just call it a day…I can't get no more positive. And that was the thing and actually I know better, but honestly I don't use condoms" p.560

"…if you're looking at giving support for the guys at the clinic, you want to cover mental health, physical health, sexuality, safe sex, drugs, and alcohol. We really need it. I'm sitting in a room with a group of guys I'm having a blast with and would like to get to know better; we're enjoying it right now…and it makes you feel good about yourself. " p.557-558

Responses from a qualitative study investigating approaches to better sexual health in HIV-positive gay men in the United States (Vanable, Carey, Brown, *et al.*, 2012).



4.1 Summary

In the previous chapters I have investigated the surveillance systems in place for LGV in the United Kingdom and the profile of those who acquire LGV. I have looked at LGV re-emergence from the point of view of individual's risk-behaviour and the response of the health care system. This chapter will focus on the wider determinants of STI re-emergence.

Seroadaptive behaviours are widely cited in the literature as strategies employed by MSM to reduce HIV transmission risk. They have also been proposed as contributing to the increasing STI diagnoses in HIV-positive MSM, and they have been hypothesised as a reason for the strong association between LGV and HIV. In order to understand the social context in which the re-emergence of LGV has appeared a critical analysis of available evidence is needed. In this chapter, I link individual, partnership-level, social and structural factors that interact with seroadaptive behaviours.

I performed a literature review on seroadaptive behaviours in HIV-positive MSM, and the information was analysed from a social epidemiological perspective which allows the incorporation of multiple intersecting levels. The arising themes were divided into structural, community, interpersonal and intrapersonal level factors.

Based on the results I will propose a conceptual framework of seroadaptive behaviours among HIV-positive MSM. In cross-sectional surveys HIV-positive MSM report seroadaptive behaviours. I show that the processes and barriers to safer sex are associated with the formation of subcultures whereby separate social rules of HIV disclosure and perception of risk depend on the setting where partners are encountered. However, these have rarely investigated whether the behaviour was part of a premeditated strategy. The clearest evidence for seroadaptive intentions comes from stable partnerships where the responsibility for partners' health motivates the use of harm reduction strategies. The sense of responsibility wanes or is shifted from the HIV-positive man to their partners when the number and anonymity of sexual encounters increases.

Overall, seroadaptive behaviours are widely adopted by MSM in high-income countries and are an important way for HIV-positive men to manage and enjoy their sexual lives. Seroadaptive behaviours are strongly context-dependent and can either reduce or increase transmission risk for infectious diseases. Future health promotion programmes should consider the setting in which HIV-positive MSM meet their sex partners and sexual health advice should be designed accordingly.

4.2 Introduction

Men who have sex with men have been disproportionately affected by the HIV epidemic. The impact of the epidemic has been high due biological and behavioural factors which affect the transmission probability depending on role versatility (MSM are able to take insertive and receptive roles in a sexual act) and sexual role positioning (HIV-transmission probability is larger for the receptive partner than for the insertive partner) (Baggaley, White & Boily, 2010; Goodreau & Golden, 2007). HIV prevalence in MSM is significantly higher than HIV prevalence in the general adult population in every setting (Beyrer, Baral, van Griensven, *et al.*, 2012). Furthermore as a result of the on-going HIV epidemic and the increase in survival due to treatment, the number of HIV-diagnosed MSM is rising steadily in the high-income countries, and the demography of MSM communities is changing as a result (Hart & Elford, 2010).

Strategies have been developed by MSM communities in response to the epidemic to reduce onward transmission. These either try to alter the mixing patterns in the population with like-with-like mixing regarding serostatus (serosorting)[21], or given a serodiscordant partnership, trying to reduce the transmission probability by the HIV-negative partner taking the insertive position in anal sex (strategic positioning) or trying to limit exposure to the virus (for example with withdrawal before ejaculation or refraining from anal sex) (Parsons, Schrimshaw, Wolitski, *et al.*, 2005). There is evidence to demonstrate HIV-positive men are more

[21] Though serosorting can mean any partnering according to HIV status, it is most commonly measured in literature as unprotected anal intercourse (UAI) between partners of same status. This is also what serosorting is understood as in this chapter.

likely to have UAI with other HIV-positive men[22] than with HIV-negative men (Crepaz, Marks, Liau, *et al.*, 2009). Seroadaptive behaviours can be seen as a harm reduction strategy (McConnell, Bragg, Shiboski, *et al.*, 2010) or a functional response to the epidemic given the awareness of HIV transmission risk factors (from Kurtz *et al.* 2012 who referred to McConnell *et al.* 2010 and Van de Ven *et al.* 2002a). In its widest definition seroadaptive behaviour can be seen as any modification of sexual behaviour depending on the person's serostatus, the (perceived) status of the partner and/or HIV transmission risk by type of sex act (Vallabhaneni et al. 2013, who referred to definition by Le Talec & Jablonsky 2008).

In Germany (Marcus, Schmidt & Hamouda, 2011) HIV-positive men who reported either strategic serosorting (preferred partners of the same status) or tactical serosorting (use condom if a partner is not the same serostatus or is of unknown status) were more likely to have had a bacterial STI within the past 12 months than non-serosorting HIV-positive men (over two-fold increase for strategic serosorters, and over three-fold increase for tactical serosorters). Neither form of serosorting was associated with bacterial STIs among HIV-negative men. Also in San Francisco the levelling off of HIV incidence at the same time as an increase in syphilis and gonorrhoea was suggested to be due to increase in serosorting (Truong, Truong, Kellogg, *et al.*, 2006). For HIV-positive men seroadaptive behaviours have the potential to reduce risk of onward HIV transmission (McConnell, Bragg, Shiboski, *et al.*, 2010) but increase the risk for acquiring and transmitting STIs.

The effect of serosorting for HIV-negative men is less clear as the proportion of undiagnosed HIV-infection in the population will affect the reliability of seemingly concordant seronegative partnerships (Wilson, Regan, Heymer, *et al.*, 2010). Thus for HIV-negative men to serosort HIV-testing becomes the primary means of HIV prevention (Eaton, Kalichman, Cain, *et al.*, 2007). In a pooled longitudinal analysis of seroadaptive behaviours in HIV-negative MSM who reported serosorting had a higher HIV infection risk than men who reported no UAI or men who reported only having one HIV-negative partner; however HIV-negative serosorting men were at lower risk than men who did not have any seroadaptive behaviour (Vallabhaneni, Li, Vittinghoff, *et al.*, 2012). For HIV-positive men seroadaptive behaviours are in place to limit onward transmission and as a way of managing one's sexual identity in the presence of HIV-infection. Therefore HIV disclosure and protected sex is

[22] HIV-positive in this chapter is used interchangeably with HIV-diagnosed.

no longer about protecting oneself but it becomes a moral responsibility and an altruistic act towards the partner.

To separate the different levels that influence sexual behaviour in HIV-positive MSM this chapter presents a conceptual framework of seroadaptive behaviours employed by HIV-positive MSM. They were chosen as the primary group of interest as high-risk HIV-positive MSM have been identified as the group most affected by the LGV re-emergence. A seroadaptive behaviour framework for HIV-negative men would likely be different, with the emphasis on strategies for preventing the acquisition of HIV instead of managing one's life in the presence of HIV.

In this chapter I used a social epidemiology perspective, which focuses on social causes of disease, with the intention of identifying mechanisms through which behaviours at the population level may be influenced. Social epidemiological frameworks have previously been used to assess factors that affect HIV transmission dynamics (Poundstone, Strathdee & Celentano, 2004) where the factors were divided into individual, social and structural with the emphasis on macro-level determinants to describe variation in disease outcomes. The development of this framework was also influenced by the social resistance framework by Factor *et al.* (2011) which both complements and challenges the social epidemiology approach of Poundstone *et al.* by describing structural inequalities whilst at the same time advocating the role of individual-agency. In Factor's framework, health disparities are explained by the non-dominant group being firstly detached from the culture of the dominant group and secondly their collective identity as a non-dominant group causing an opposition against the (health) values created by the dominant group. This approach is interesting given the minority status of MSM in the predominantly heterosexual world, and the additional minority status HIV-positivity brings along. Factor's resistance framework also enables us to contrast the personal responsibility of an HIV-positive individual and the wider determinants for his health and behaviour.

Previous literature has focused on conceptual frameworks for understanding reasons for increases in STIs in MSM (Fenton & Imrie, 2005). A systematic review and narrative analysis on barebacking (intentional unprotected anal intercourse) was done by Berg (2009) to explore the different macro and micro-level factors

associated with bareback sex. Crepaz *et al.* (2009) performed a meta-analysis of prevalence of unprotected anal intercourse in HIV-diagnosed MSM, and they also indirectly measured serosorting and strategic positioning. However to my knowledge no study to date has synthesised the literature on seroadaptive behaviour in HIV-diagnosed men from a social epidemiological point of view.

Reviews, in particular systematic reviews, allow for a more objective assessment of available evidence (Egger, Davey Smith & Altman, 2001). A systematic literature review addresses a well-defined research question and uses a pre-specified search strategy and selection criteria for studies allowing for transparency in information retrieval and analysis.

There is a lack of precise definitions for seroadaptive behaviours, and this study tries to understand the context in which these behaviours take place. This can offer a possible explanation for the re-emergence of a once rare STI like LGV. However, due to the scope of this exercise, the study is more appropriately called a broad literature review rather than a systematic review. For the same reasons, I decided it would be premature to quantitatively estimate the prevalence of seroadaptive behaviours and chose a narrative synthesis instead.

4.3 Aims

The overall aim of this chapter is to better understand the social context in which seroadaptive behaviours occur. A literature review is used to develop a social epidemiological framework of factors contributing to seroadaptive behaviours and beliefs. Three specific objectives were defined:

- To perform a literature review on seroadaptive behaviours in HIV-diagnosed MSM
- To synthesise the findings.
- To propose a conceptual framework based on the literature.

4.4 Methods

The aims were addressed by an extensive review and narrative synthesis of literature on seroadaptive behaviours. Peer-reviewed articles were reviewed if they presented results of seroadaptive strategies / behaviours in HIV-diagnosed MSM.

The following search words and Boolean operators were used: serosorting OR seroadaptive OR "strategic positioning" OR "sexual harm reduction" OR serodisclosure OR (serostatus AND disclosure) OR "negotiated safety" OR serodiscordan* OR seroconcordan* OR ("hiv status" AND partner) AND ("men who have sex with men" OR homosexual*). PubMed and Web of Knowledge were used as the search platforms, the first covers Medline and the latter the following databases: CAB Abstracts, Conference Proceedings Citation Index, Derwent Innovations Index, ISI Web of Science, ISI Current Contents, ISI Journal Citation Reports and Scientific WebPlus.

The search results were managed, merged and checked for duplicates in EndNote and Mendeley citation managers, and a review database was created in Microsoft Excel. Reference lists of main articles were investigated to supplement the search. The titles were first screened and titles not covering HIV or MSM were excluded. Then abstracts were scanned and the inclusion criterion was that the article focused on HIV-positive MSM or indicated that results were stratified. Articles of HIV-negative men, HIV testing or HIV seroconversion were excluded. After this the articles were screened, and all articles where results were stratified for HIV-positive MSM with some information of sexual risk taking and partner's serostatus were included (also articles which addressed HIV disclosure to sex partners were included regardless of whether they measured sexual risk taking). In addition new articles on MSM were browsed through to identify the most recent publications. Given the context of the thesis and LGV emergence in UK, additional search for seroadaptive behaviours and beliefs was done on Sigma Research (http://www.sigmaresearch.org.uk/) as some of their reports would not have been published in peer-reviewed journals. No criteria were predefined for study design. Authors of the studies were not contacted during the review.

Four categories were predetermined as levels of the framework: structural factors, community level factors, interpersonal factors, and intrapersonal factors. This was adapted from the structure of Poundstone's (2004) social epidemiological framework for HIV. Articles included in the review were first categorised according to their perceived framework level, and reviewed accordingly. In the narrative synthesis, studies where seroadaptive behaviours were a main component and studies focusing on HIV-diagnosed men and their partnerships were given more attention. After each section, a summary table is presented which shows the main themes that emerged from the literature, how they are potentially related to seroadaptive behaviours and how I hypothesise these to be related to HIV and LGV transmission risk.

4.5 Results

4.5.1 Literature review

The literature review process is presented in Figure 1. Searching the databases resulted in 633 articles. Screening of titles lead to exclusion of 17 articles, and after reading the abstracts further 272 articles were excluded (they were mainly articles which focused on HIV-negative MSM, HIV testing rates, seroconversion risk factors, the results were not stratified by HIV status or the focus group was not MSM). After that 344 articles were screened; 39 of which I was unable to find or access. 120 articles were screened and excluded (the results were not stratified for the HIV-positive MSM respondent, or there was no information on the sex partners' perceived serostatus). Thirteen additional articles were included through reference lists and through Sigma research. This led to the inclusion of 199 articles which I reviewed for this study. The majority of the articles were cross-sectional studies which often presented an aggregate summary table of respondents' partnership characteristics and/or partner number. Also longitudinal and qualitative studies, meta-analyses and broad literature reviews and a small number of intervention-based studies were identified.

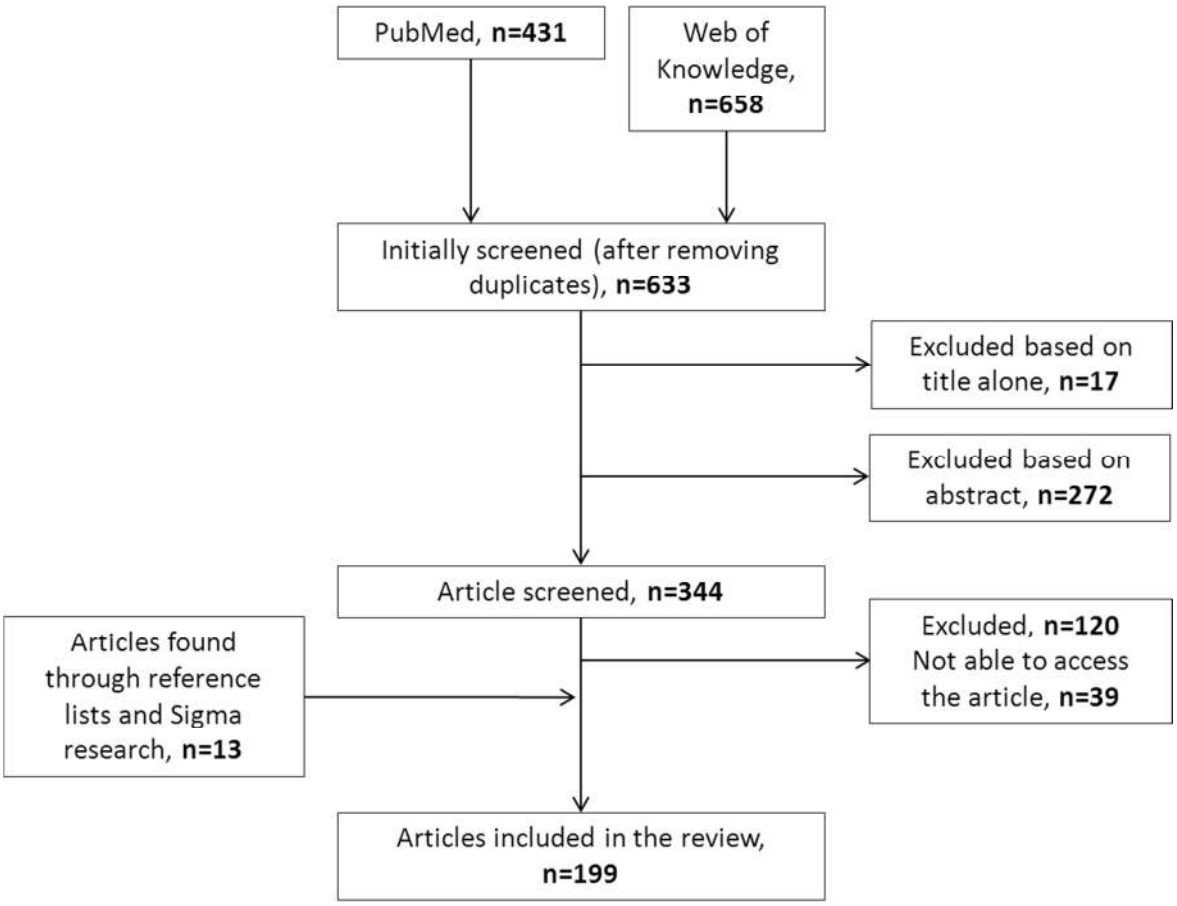


Figure 1. Inclusion and exclusion of articles in the review.

Figure legend: The flowchart describes the process of literature review. It should be noted that not all 199 included in the review are cited in this chapter. A full list of the 199 articles is in the appendix for chapter 2.

The literature review demonstrated that there are a number of prerequisites for seroadaptive behaviours to occur: it requires awareness of HIV transmission risks, availability and uptake of HIV testing, disclosure of HIV status, sufficient prevalence of HIV in the community and HIV-related attitudes in the community that facilitate seroadaptive behaviours (such as sexual market places which accommodate or specifically cater for HIV-positive men). In the next sections I will go through seroadaptive behaviours from structural, community, interpersonal and individual-level perspective.

4.5.2 Structural factors

4.5.2.1 Geographical region and culture

The majority of studies identified in this review were conducted in the United States, with San Francisco, New York and Seattle being the most represented cities. Seroadaptive research papers have also been published from Australia, United Kingdom, Netherlands, Germany, France and Canada.

Few studies have been conducted outside high-income countries: in a study of MSM in Kuala Lumpur among whom prevalence was low (3.9%, 20/517) they found that the participants had a great deal of wrong information on transmission routes (with around a fifth believing HIV is not transmitted through anal sex) (Kanter, Koh, Razali, *et al.*, 2011). The lack of information in itself would hinder employing seroadaptive strategies. In an online survey conducted in Asia (Wei, Lim, Guadamuz, *et al.*, 2012), there was variation in disclosure patterns for HIV-positive MSM between countries: Chinese men were the least likely to disclose (88% non-disclosers) followed by Japanese (74%) whilst Philippine men were more likely to disclose (53% non-disclosers), although sample size per country was relatively small with only 416/13,883 of the total sample self-identifying as HIV-positive. Not knowing partner's HIV status was the strongest explanatory factor for non-disclosure in multivariate model. The authors interpret this as a sign of serosorting being less common in Asian populations, and attribute it partially to stigma which inhibits disclosure. This interpretation is supported by a study from Bangkok, Thailand, which looked at disclosure among HIV-positive MSM (Edwards-Jackson, Phanuphak, Van Tieu, *et al.*, 2012): disclosure (to the last sexual partner) was rare, with 60.5% reporting no disclosure; however, only 17.2% of men reported UAI. In another Bangkok-based study, the HIV status of the steady male partner of an HIV-infected MSM was not associated with having reported unprotected sex in the past 3 months (Sirivongrangson, Lolekha, Charoenwatanachokchai, *et al.*, 2012).

4.5.2.2 Criminalisation of HIV

Criminalisation of a type of behaviour is an extreme form of control over the behaviour of the population, and criminalisation of HIV transmission might influence how HIV-positive individuals choose their partners. In England and Wales there have been criminal prosecutions for recklessly transmitting HIV and other STIs (reckless cause of serious bodily harm, Section 20 of the Offences Against The Person Act 1861) (Dodds,

Bourne & Weait, 2009)[23] and the first conviction of sexual transmission of HIV in the UK happened in 2001 in Glasgow (Chalmers, 2002).

In sexual partnerships it is debatable as to what constitutes consent, what are the levels of risk for HIV exposure and how the responsibility of risk should be shared if one of the partners is HIV-infected. Chalmers (2002) argues that criminal law has a role in shaping attitudes and modifying behaviour and in a qualitative study of HIV-positive men in Seattle and Los Angeles, men who always disclosed their status expressed fear over legal consequences of not disclosing as one of the reasons they disclosed (Gorbach, Galea, Amani, *et al.*, 2004). On the other hand qualitative research by Dodds *et al.* (2009) showed the challenges of attempting to influence behaviour: aside from the misconception surrounding which situations could be prosecuted for[24], around half of the 29 men who discussed the personal impact of the laws said they had at least planned to change their behaviour in response to legal concerns. The others said they were not influenced by the law, and they reasoned this by stating their sexual activities were not such that they were likely to be prosecuted; one respondent stated he had safe sex unless he knew the other person was HIV-positive too. Five of the men interviewed took another perspective and explained that they had increased the level of anonymity in their sex lives and reduced their openness about being HIV-positive, for instance by changing information on their online profiles. In general, men with higher numbers of partners found anonymous sexual settings less risky for legal consequences.

An example of legal ramifications comes from a case study where the Ottawa Police in Canada released a photo, name and sexual orientation of a person who had continuously refused to disclose his HIV status in situations that were described as posing risk of HIV transmission. A study by O'Byrne *et al.*, (2013) aimed to look at HIV testing rates post the media uproar on the case and performed a qualitative study among MSM in Ottawa. They found no significant changes in the overall pattern of HIV testing and the qualitative interviews

[23] However as reviewed by the Sigma research, there has been a decline in successful prosecutions as timing and direction of transmission is difficult to establish if the defendant does not plead guilty (Dodds, Weatherburn, Bourne, *et al.*, 2009).

[24] Disclosure is a pre-condition for consent, but legally one requires the susceptible individual to understand and consent to "*the risks associated with unprotected intercourse with a person they know has HIV*", p.139 in Dodds, Bourne & Weait, (2009)

revealed a mixture of condemning views of HIV-positive people who have unprotected sex (labelling them as criminals or murderers) and concerns about how HIV-positivity automatically changes the authority's perspective on you. The connectedness of local public health officials and the police force arose as a theme from the interviews with some participants being concerned about the records that were being kept of them and mentioned that anonymous HIV testing might be a safer option in the future given the potential for legal charges. When MSM in the United Kingdom were asked about their opinions on criminal prosecutions responsibility was one of the main themes discussed (regardless of whether the men were in favour of criminal prosecutions or not).The intention of the HIV-infected person, the potential harm caused and the benefit or harm of imprisoning "HIV transmitters" also emerged from the study (Dodds, Weatherburn, Bourne, *et al.*, 2009). Based on the report it seemed as if the proximity of risk for HIV acquisition and agreement with criminal prosecutions were inversely related (with younger men, bisexuals and those with fewer partners being more in favour of the prosecutions whilst London-based older men with higher number of partners were more likely to disagree with the prosecutions).

In the general practice setting, legal obligations of clinicians are not always clear either: four hypothetical situations were presented to Irish physicians of a newly diagnosed HIV-positive patient in a hetero- or homosexual partnership of two years. The patient does not intend to tell the partner of his infection and either protected or unprotected sex is indicated. The clinician was more likely to break patient confidentiality if the relationship was heterosexual (than homosexual) and if unprotected sex was reported. Clinicians who had broken patient confidentiality before and who had less years in the profession were more likely to favour informing the patient's partner at the expense of patient confidentiality (Daly, Hevey & Regan, 2011).

4.5.2.3 Ethnicity, stigma, poverty and homophobia/heterosexism

In a meta-analysis Millett *et al.* (2007) explore the variation in risk for HIV infection for black MSM in North America and the United Kingdom. In a sub-analysis for HIV-positive men, black men were less likely to disclose their HIV status to partners when compared to other HIV-positive MSM (summary estimate from three studies: OR 0.46; 95% CI 0.26, 0.77). However they were not statistically different in their reporting of serodiscordant UAI, serosorting or strategic positioning (when measured across several studies). There were inequalities regarding access to HIV care, with black men having lower CD4 cell counts, being less likely to be

on ART, having poorer ART adherence and with fewer achieving viral suppression. In addition, the study revealed structural inequalities with more black MSM on low income or with low education. Another study, based in San Francisco, found no differences in seroadaptive strategies used by Asian and Pacific Islanders, white, black, Latino or other ethnic group though black and Latino men were less likely to be consistent in their risk-taking (Wei, Raymond, Guadamuz, *et al.*, 2011).

In a study by Bachmann et al. (2009) white HIV-positive men reported more sex partners than black HIV-positive men. Black men also reported fewer HIV-negative and unknown HIV status partners than white men. Among focus groups of African American non-gay identified MSM HIV-related stigma, homophobia, racism, low socioeconomic status, substance abuse, and low priority given to one's health were mentioned as issues surrounding condom use, disclosure and sexual activity (Harawa, Williams, Ramamurthi, *et al.*, 2006).

The impact of ethnicity on gay men's lives has been mainly explored in the United States, and in general African American MSM have been reported to discuss their HIV status less with their partners than white MSM or Latino MSM. (Bird, Fingerhut & McKirnan, 2011; O'Leary, Fisher, Purcell, *et al.*, 2007; Overstreet, Earnshaw, Kalichman, *et al.*, 2012). Comfort with one's sexual orientation and social capital have been associated with greater disclosure and communication of HIV status in Latino men (Zea, Reisen, Poppen, *et al.*, 2003).

Bird and Voisin (2011) present a conceptual model of factors contributing to disclosure decisions of HIV status with casual sex partners. The model is based on HIV related stigma and how this is the primary component affecting disclosure. Stigma is mediated through beliefs about disclosure risks, beliefs about responsibility to disclose (as opposed to right to privacy regarding one's health), assumed partner characteristics (such as HIV status and risk perceptions), the setting of the encounter and how much communication is assumed, and also by managing one's anxiety by looking for information which confirms your assumptions (non-disclosers can be assumed to be of the same status as oneself).

In another theoretical model that was validated with data, Johnson *et al.* (2008) found that internalised heterosexism (negative attitudes towards homosexuality) was indirectly associated with greater HIV-

transmission risk (UAI with HIV-negative or unknown partner) and poorer ART adherence. The association was mediated through negative emotional states and more regular stimulant use. In a data-analysis by Rendina *et al.* (2012) sexual compulsivity was also associated with more HIV-related stigma.

Table 1. Summary of structural factors contributing to seroadaptive behaviours in HIV-positive MSM, and their hypothesised association with transmission risk.

Structural factors	Potential relation to seroadaptive behaviours (SABs) in HIV-positive MSM	Hypothetical relation to HIV/LGV transmission risk
Geographical region	Most evidence of SABs originates from high-income countries.	Risk in relation to SABs is likely to vary between regions.
Criminalisation of HIV	Behaviour may be altered due to concern over prosecution.	Uncertain how criminalisation affects behaviour and transmission risk.
HIV-related stigma	Affects disclosure of HIV status	Likely to increase transmission risk for HIV, and therefore also for LGV.
Hetero-normative environment	Associated with sdUAI	Likely to increase transmission risk for HIV, and therefore also for LGV.
Ethnicity	Black men are less likely to disclose their HIV-status, but few behavioural associations are seen to indicate systematic differences in SABs	Relation to transmission risk not clear. The way ethnicity is linked to other structural barriers might be more influential on transmission risk than SABs.

4.5.3 Community factors

4.5.3.1 Gay community social norms

Community perception of HIV can be an important factor on determining norms around sexual behaviour, stigma and disclosure. Flowers et al (2000) examined the change in attitudes towards who carries the responsibility for HIV, and discussed how gay men were blamed for the HIV epidemic and were socially separated from the "general population" as a specific risk group. The authors go on to describe how, at the early stages of the epidemic, gay community advocates promoted the idea that everyone, including oneself, should be assumed to be infected which created a shared accountability for HIV transmission, where the risk and responsibility was managed equally by everyone in the community. As HIV antibody testing became more

widely available the community risk became an individual or partner-level risk, and a "hierarchy of risk" emerged where groups were distinguished as being at risk (HIV-negative), potentially posing a risk (HIV untested) and posing a risk (HIV-diagnosed). HIV testing has increased this "othering", as they call it, by marginalising the HIV-positive men (Flowers, Duncan & Frankis, 2000).

In a qualitative study in Scotland (Davis & Flowers, 2011) in a low HIV prevalence area, described by the authors as lacking an HIV-positive subculture, men discuss the imbalance of a serodiscordant relationship with HIV-positive men having additional responsibility in the relationship. HIV-negative men are in a more vulnerable position but this also gives them indirect power and some described unsafe sex as a gift they can give to their HIV-positive partner. In general as the partnership duration grew longer the rules around safe sex were described to weaken, and the fear of acquiring HIV and its consequences were perceived as less severe.

In an earlier qualitative study (Flowers, Duncan & Frankis, 2000), also from Scotland, similar attitudes on HIV-related stigma were identified where men who were known to be infected were described (by HIV-negative and –positive men) as having the responsibility to disclose, to manage the exposure and to take the blame if transmission occurs. The authors described the setting as having a "universal HIV-negativity assumption". The social cost of testing HIV-positive was also given as a reason not to get tested as untested men were equivalent to HIV-negative men in terms of responsibility for sexual risk taking. At the other end of the spectrum HIV-positivity can sometimes be seen as the norm as was the case in another qualitative study based in San Francisco (Sheon & Crosby, 2004). The study describes how an HIV-positive subculture dominates with a reverse stigma and HIV-negative men feel excluded as everyone is assumed to have HIV or not be concerned over the infection. Apathy towards disclosure and unprotected sex was discussed by the men in the study and in some occasions men did not want to disclose their negative status to a man perceived to be positive in fear of rejection.

Among MSM from Toronto, sampled for preferring UAI, the participants talked about "unspoken knowledge" regarding HIV which reduces need for formal disclosure (Adam, Husbands, Murray, *et al.*, 2008). An example of this was internet profiles where HIV status was asked for. According to one participant, leaving this blank

indicated HIV-positivity as anyone HIV-negative would have no reason not to declare their status. Another man gave an example of profile stating "safe sex only" being interpretable as the person having HIV. Furthermore it was generally perceived that bathhouses were environments where different rules applied regarding need for disclosure. How the underlying norms affect sexual behaviour over time is not clear and different locations and survey methods report differing risk behaviour trends: in Seattle no significant difference was seen in behaviours reported by STD clinic attendees between 2000 and 2006 (Menza, Kerani, Handsfield, *et al.*, 2011) whilst a study from London reported an increase in both seroconcordant and serodiscordant UAI in HIV-positive men between 1998 and 2008 (Lattimore, Thornton, Delpech, *et al.*, 2011b).

Gay-specific venues and locations are an important aspect of the gay community structure and form a social setting for meeting people. Some of these settings, such as internet and sex-on-premises venues, are also favourable environments for high-risk sex. These are explored in more detail in the next sections.

4.5.3.2 Internet

A plethora of research exists of gay men's internet use with a consensus that online sex seeking is associated with increased risk behaviour: in a meta-analytic review of online risk behaviour in MSM, Liau *et al.* (2006) identified two UK-based studies (Bolding, Davis, Hart, *et al.*, 2005; Elford, Bolding & Sherr, 2001) where HIV-positive serosorting was more common in men who looked for sex partners online than in those who did not (32-38% versus 4-8%); however, serodiscordant UAI (sdUAI) prevalence was also higher in those who sought sex online. Berry *et al.* (2008) investigated whether the internet facilitates serosorting; for HIV-positive men seroconcordant partnerships were somewhat more common when formed online than in bars or dance clubs (aOR 1.6; 0.9, 2.7) but this was not a significant difference. No difference was found for serodiscordant UAI either. In theory the internet facilitates disclosure of HIV status prior partnership formation but in a study by Horvath et al. (2008) 50% of HIV+ men reported not stating their HIV status in any of their internet profiles (with 25% stating they are HIV-negative in some or all profiles), and only 27% reported always declaring their HIV-positivity in their profiles. HIV-positive respondents were also more likely to state they were interested in unprotected sex and less likely to state they wanted safer sex.

In a large French study by Leobon *et al.* (2011) which recruited almost 14,000 MSM using gay websites (both mainstream websites as well as websites specialised into unprotected and fetish sex), 63% of HIV-positive men entered the study via a specialised gay website compared to 32% of HIV-negative men. HIV-positive men were also more likely to report sensation-seeking and depression than HIV-negative men but they were also more likely to report having mixed with the gay community in the past year. A wider range of sexual practices and higher sexual activity levels were reported by the HIV-positive men and they had more UAI in all of the reported partner serostatus categories[25].

Based on another French study HIV-positive men who reported serosorting with a casual partner were more likely to look for sex partners online (aOR 2.16; 95% CI 1.00,4.67; p-value 0.051) and less likely to look for partners in cruising venues (aOR 0.28; 95% CI 0.13-0.60) when compared to HIV-positive men who reported UAI with a casual partner but not serosorting (Velter, Bouyssou-Michel, Arnaud, *et al.*, 2009).

However as Sheon and Crosby (2004) argue in their discussion, "barebacking as a social movement" is partly facilitated by the internet, but to which extent internet has modified perceptions of unprotected sex and HIV disclosure is to be researched.

4.5.3.3 Venue type

Xia *et al.* (2006) looked at how different sexual risk behaviours and HIV prevalence are affected by sampling different venue types in California. They found that sdUAI was most common in circuit parties (30%) with also the highest HIV prevalence amongst attendees (41%). Least sdUAI was reported by those who did not visit a gay venue (4%) with HIV prevalence of 20%. In another study public sex areas were described as the territory of HIV-positive men with public sex "rules" of no disclosure, no condoms (ins, Reback, Shoptaw, *et al.*, 2005). Among men attending Californian circuit parties (Patel, Taylor, Montoya, *et al.*, 2006) HIV-negative men were more likely to prefer a partner of the same serostatus than HIV-positive men (88% versus 44% of the respondents).

[25] Same, unknown, different, or did not care about the serostatus

In a qualitative study 40 men (of whom 20 were HIV-positive) who had attended or hosted a sex party were interviewed and various themes emerged (Mimiaga, Reisner, Bland, *et al.*, 2010): sex parties were an environment for easy sex while offering a more intimate setting than a bathhouse as relatively few people attend at one time (sex parties often take place in someone's home, hotels or backroom of a club/bar). Therefore it was described as a safe place for anonymous sex, and although HIV disclosure is rare ("*it's pretty much an unspoken thing*" HIV-positive man, p.667) the hosts reported seeing HIV risk reduction practices taking place. Substance use was prevalent, and amongst some, the main reason to attend the parties. It was also reported that risky behaviour and injuries were more likely to occur when drugs were taken. Respondents also talked about the social norms and cliquey environment surrounding sex parties with some hosts being more exclusive about the attendance than others.

Another highly selective setting is the POZ parties which were created as events to facilitate sex between HIV-positive MSM; they started in mid-1990 in New York City, and the concept has subsequently been exported to other US cities and Europe. The most important reason (amongst 81 respondents for a study done by Clatts e*t al.* (2005)) for attending a POZ party was "Not having to worry about HIV disclosure" with "like having uninhibited or unrestricted sex" as the second most important followed by "don't have to worry about infecting others".

Table 2. Summary of community-level factors contributing to seroadaptive behaviours in HIV-positive MSM, and their hypothesised association with transmission risk.

Community-level factors	Potential relation to seroadaptive behaviours (SABs) in HIV-positive MSM	Hypothetical relation to HIV/LGV transmission risk
Social norms	Depending on setting the norms can either support or marginalise the HIV-positive population.	Can either increase or decrease transmission risk.
HIV prevalence	Sufficient pool of HIV-positive (who are diagnosed) individuals needed for SABs to take place. May also affect the formation of HIV-positive subcultures.	Increasing prevalence is likely to increase HIV transmission risk (unless SABs were consistently practiced).
Internet	Facilitates SABs through specialised websites.	Good evidence for increased risk for HIV and LGV.
Sex-on-premises venues	Unless the venue/event is for HIV-positives, increased anonymity makes consistent SABs less likely.	Good evidence for increased risk for HIV and LGV.
Social networks	This factor is related to all the factors above. Social and sexual networks are likely to have some overlap, and attitudes in one's social network are likely to have an influence on sexual behaviours.	Can either increase of decrease transmission risk.

4.5.4 Interpersonal factors

Given seroadaptive behaviours occur within partnerships, a large proportion of identified articles focused on partnership-level factors. Definitions of seroadaptive behaviours have arisen later, but the practices were reported from early on in the epidemic; in a sample of HIV-positive MSM from Los Angeles in 1991 it was noted that though less than 10% of the HIV-positive participants reported unprotected sex, when it happened it seemed to occur more with HIV-positive men (Marks, Ruiz, Richardson, *et al.*, 1994) and seroconcordant partnerships were more likely to report UAI than discordant partners in 1992 (Hoff, Stall, Paul, *et al.*, 1997). A qualitative study from the UK, between 1995-1996, describes one man's relief of finding someone who is also HIV-positive: "*I met this lad and we went back to his flat and I told him I was HIV and he said he was as well so*

it was like, `Wow'. It was a wonderful feeling, I know it shouldn't' have been but at the time I was bowled over. It was just a normal sexual relationship." (p.55 in Holt et al. 1998).

In the following sections I will cover the interplay between HIV disclosure and partner and partnership characteristics.

4.5.4.1 HIV disclosure

Poppen et al. (2005) hypothesise that HIV disclosure, seroconcordance and type of relationship between partners act jointly creating a context where unprotected sex can occur, and they demonstrate that seroconcordance predicts well UAI, possibly as a result of decision making process between partners. The authors discuss that though seroconcordance may be a primary determinant for UAI, it is correlated with relationship as disclosure is more likely to occur in long-term relationships.

However, disclosure is not a binary construct, and those HIV-positive men who were labelled as inconsistent disclosers reported more high risk behaviours than non-disclosers or consistent disclosers[26] (Parsons, Schrimshaw, Bimbi, *et al.*, 2005). Self-efficacy, intention and connection to other HIV-positive men were associated with being a consistent discloser in this study, and it was suggested that inconsistent disclosers lacked strategies to manage their sexual risk taking whilst non-disclosers may use alternative ways to avoid high risk sex. In an Australian study (Holt, Rawstorne, Worth, *et al.*, 2011) HIV-positive "disclosers" were compared to HIV-positive "non-disclosers" (to last casual UAI partner) and the authors found that HIV disclosure is more common if the sex has occurred between partners before, if the respondent believes that HIV-positive men have the responsibility to disclose before sex and if the respondent has sought health information about HIV. In an earlier study of a sample collected from Los Angeles in 1995 safer sex without disclosure occurred more often if the partner's HIV status was unknown or there was less of an emotional bond with the partner (Marks & Crepaz, 2001).

[26] inconsistent disclosers disclosed to some "non-main" partners, non-disclosers disclosed to none of their non-main partners and consistent disclosers disclosed to all of their non-main partners

Similar findings were reported by Gorbach *et al.* (2004) from their qualitative interviews with men from Seattle and Los Angeles where non-disclosure was reported with being in denial about one's HIV-positivity, thinking your medical condition was nobody else's business, having a low viral load, fearing rejection, casual encounters, drug use and public places for sex. Disclosure depended on type of sex had (whether condom was used, and whether the sex act posed a high or low transmission risk), partner asking about status or the partner disclosing first, having feelings for the partner, sense of responsibility over transmission and a fear of legal consequences of non-disclosure.

The idea of seroguessing is introduced in a study where both HIV-positive and –negative MSM were more likely to have UAI if they knew or guessed the HIV status of their partner (Zablotska, Imrie, Prestage, *et al.*, 2009), and another study found that a perceived positive serostatus was based on behaviour, normative assumptions (location where partner met) and circumstantial evidence (such as mentioning medication) (Parsons, Severino, Nanin, *et al.*, 2006).

4.5.4.2 Partners and partnerships

Relf, Bishop, Lachat, *et al.*, (2009) found three different concepts of sexual relationships among urban American men with HIV: dissonance to sex (avoiding sex entirely), having "just sex" or being in a sexual relationship that was "going somewhere". Sexual behaviours and disclosure were strongly dependent on the type of sexual relationship the men reported. The least comfortable to disclose were those who avoided sex whilst having encounters "just for sex" were coupled with less obligation to discuss HIV positivity. In these instances lack of disclosure was managed by practicing safer sex (condom use or sexual practices that are less risky than UAI) unless the other person had come forward as HIV-positive. For more stable relationships disclosure and individual responsibility was seen as important and respondents in these circumstances reported protected sex with HIV-negative partners, whilst with HIV-positive partners safer sex was practiced but definitions of safer sex were less strict than in serodiscordant partnerships. In interviews with HIV-positive and HIV-negative MSM Frost et al. (2008) found that the men were concerned of acquisition or transmission of HIV and saw serodiscordant sex as inherently posing a risk. Intimacy in a seroconcordant partnership was reported as an important contributing factor to seeking such relationships. HIV status was important but not the most important factor in men's perceptions of successful partnerships.

A study by Theodore *et al.* (2004) reported intimacy in partnership increased condom use, but both seroconcordant and serodiscordant couples stated intimacy as a reason to have UAI. However drug use before or during sex made condom use less likely irrespective of levels of intimacy in the partnership. In another study intimacy in partnership was associated with UAI in the partnership (as were level of attachment, feeling lonely and support in the partnership) (Hoff, Colleen, Chakravarty, *et al.*, 2012).

4.5.4.3 Type of sex and number of sex partners

In a meta-analysis of UAI among HIV-positive MSM (Crepaz, Marks, Liau, *et al.*, 2009) the prevalence of receptive UAI was higher than insertive UAI when HIV-positive men had HIV-negative partners (5%, 95% CI 4,7% versus 9%, 95% CI 7, 12%) which the authors interpret as an indication of strategic positioning but note that as the studies are cross-sectional in nature and give an aggregate measure, we cannot directly deduce temporal relationship between HIV-status and positioning preference nor premeditated intention for strategic positioning from these results. Overall in this review I found a lack of studies that had focused on strategic positioning, whilst many studies interpreted the findings to indicate strategic positioning to occur (Parsons, Schrimshaw, Wolitski, *et al.*, 2005; Van de Ven, Kippax, Crawford, *et al.*, 2002; Crepaz, Marks, Liau, *et al.*, 2009). In a Swiss study (Dubois-Arber, Jeannin, Lociciro, *et al.*, 2012) the researchers addressed strategic positioning as a method to avoid HIV transmission directly in their questionnaire. In the responses 8% of HIV-positive men reported strategic positioning (serosorting was reported by 41% and withdrawal before ejaculation by 33% of HIV-positive men who had had UAI with casual male partners in the prior 12 months).

Among HIV-positive MSM from the UK, in the Relative safety II report by Sigma research (Bourne, Dodds, Keogh, *et al.*, 2009), many preferred UAI and reported larger partner numbers. They described the risk-management strategies such as selecting same status partners for UAI. Conversely those HIV-positive men with fewer partners rejected the idea of serosorting and wanted to distance themselves from those who sought unprotected sex. The authors of the study were surprised by how "*disgusted"* some HIV-positive respondents were of the idea of choosing HIV-positive partners, and they discuss the role HIV-related stigma and how it contributes to their dislike of the whole idea of the HIV-positive culture (with "*unrestrained sexuality*" and high incidence of STIs) and men in that setting.

Table 3. Summary of interpersonal factors contributing to seroadaptive behaviours in HIV-positive MSM, and their hypothesised association with transmission risk.

Interpersonal factors	Potential relation to seroadaptive behaviours (SABs) in HIV-positive MSM	Hypothetical relation to HIV/LGV transmission risk
HIV disclosure	Necessary precursor for informed SABs. Seroconcordancy is associated with UAI. HIV disclosure is probably related to all other factors.	If results in seroconcordant unprotected sex, can decrease HIV transmission risk, but increase LGV transmission risk. Seroguessing likely to reduce the efficacy of SAB as a harm reduction strategy.
Partner characteristics	In long-term partnerships SABs were more consistent than in short-term partnerships.	Likely to reduce transmission risk for main partner, likely to increase risk for casual partners.
Type of sex	There is some indication of strategic positioning practiced among HIV-positive MSM. Sometimes lack of disclosure was managed by choosing less risky types of sex.	Potential to reduce HIV transmission, impact on LGV transmission unclear.
Number of partners	Increased anonymity and decreased responsibility associated with number of partners, which make SABs less likely.	Potential for increased transmission risk if safer sex does not compensate for the increased anonymity.
Intimacy and support	In long-term partnerships, UAI may be favoured in place of protected sex. Concordant partnerships can increase intimacy experienced.	Increased responsibility is likely to reduce transmission risk if it translates into safer sex, but it might also increase UAI in long-term partnerships, which may increase LGV risk.

4.5.5 Intrapersonal factors

4.5.5.1 Time since HIV diagnosis

An interesting theme that emerged from longitudinal studies was the changing risk profile after HIV diagnosis. In a San Francisco based cohort study Vallabhaneni *et al.* (2013) recruited MSM in acute /early HIV infection and followed them over time. Marked changes in both partner numbers and sdUAI was reported over time with mean of 4.2 (95% CI 2.7, 6.6) potentially sdUAI partners from 3 months prior to diagnosis, which had reduced to a mean of 0.9 (95% CI 0.5, 1.7) (potentially serodiscordant UAI partners in past 3 months) by twelve months and by 48 months to a mean of 1.7 (95% CI 0.9, 3.1) per three months and by 60 months to a mean of

1.0 (95% CI 0.5, 1.9). A study looking at risk behaviour in MSM during the year post-diagnosis found a decline in reported partner numbers from 8.81 to 5.84 per three months ($p<0.0001$) with the sharpest decrease in the first 3 months. The proportion reporting HIV infected partners increased over the year (from 13.7 to 38.6%); HIV-positive UAI partners increased over the year post-diagnosis whilst of those reporting UAI none reported HIV-negative UAI partners by 12 months post diagnosis (Gorbach, Weiss, Jeffries, *et al.*, 2011). In another paper based on the same study UAI was associated with partner serostatus and partner being your main partner at follow-up (approximately 3 months from estimated infection) (Gorbach, Drumright, Daar, *et al.*, 2006). In a cross-sectional study of disclosure in Australian MSM, men who had known they were HIV-positive for more than 5 years was positively associated with disclosure to casual partners (Klitzman, Exner, Correale, *et al.*, 2007).

In Relative safety II (Bourne, Dodds, Keogh, *et al.*, 2009) the qualitative nature of changing attitudes post HIV-diagnosis is described: for those who were recently diagnosed the concern of onward transmission caused many to abstain from sex for some time whilst men who had lived with the HIV diagnosis for longer had come to terms with HIV being part of their sex lives with risk management becoming a part of it. Some described a fear of reverting back to behavioural patterns they had prior to diagnosis given the availability of bareback sex and difficulty in managing their decision of decreased sexual risk.

4.5.5.2 Intention and self-efficacy

Relatively few studies have focused on intention for seroadaptive behaviours, and most evidence comes from qualitative or longitudinal studies. In a longitudinal study from San Francisco (McFarland, Chen, Nguyen, *et al.*, 2012) condom serosorting (UAI only if partner is seroconcordant) was the most reported intention by HIV-positive men (56%) followed by withdrawal (46%), 100% condom use (42%), pure serosorting (only seroconcordant partners with whom UAI is preferred, 37%) and strategic positioning (41%). The strongest retention to intention was to not have any sex in 12 months (63% of those who had this strategy) whilst those who had sex strategic positioning (always receptive) was most likely to be adhered to during follow-up (41% who had this strategy). 20% of those who intended to adhere to 100% condom use did so. In their earlier paper (McFarland, Chen, Raymond, *et al.*, 2011) the authors found that 46.7% of men consistently employed a seroadaptive behaviour with strategic positioning being more common than pure serosorting (21% compared

to 16%). However a study from Australia reports consistent condom use being the most common HIV prevention strategy (for both HIV-positive and –negative) with almost a quarter of seropositive men reporting consistent condom use in the past 6 months and the reported seroadaptive practices were consistent which the authors interpret as intentional behaviour; however seropositioning and withdrawal before ejaculation were reported by a minority of men (Mao, Kippax, Holt, *et al.*, 2011).

Intentions for safer sex had the strongest preventative association against unsafe sex behaviour (defined as sdUAI) in a study by Miner *et al.* (2009). Personal beliefs about consequences of unprotected sex are important (Halkitis, Green, Remien, *et al.*, 2005): there was an inverse association between beliefs that HIV-positive man can re-infect another HIV-positive man or infect them with an opportunistic infection and having seroconcordant UAI whilst endorsing the idea of "STIs are not a big deal" had a positive association. In the same study seroconcordant UAI was associated with higher levels of hedonistic expectations and sexual compulsivity and less perceived responsibility of sexual partner.

High levels of HIV prevention altruism were negatively correlated with any anal sex in a group of urban American MSM and it was also inversely associated with sdUAI though the association did not remain significant in multivariate model after adjusting for methamphetamine and Viagra use and compulsive sexual behaviour (O'Dell, Rosser, Miner, *et al.*, 2008). In another study there was no difference between self-esteem measured in HIV-positive and -negative men, but there was an interaction between HIV status and self-esteem when it was used to predict self-disclosure (Moskowitz & Seal, 2011).

4.5.5.3 Risk behaviour

In a Sigma Research report (based on Gay Men's Sex Survey from 2000) those who had a high number of sex partners (30 or more in the past year) in the past year were more likely to have acquired HIV in the previous year (0.7%) and more HIV-positive men were more likely to report high number of partners (31.1%) (Davies, Reid & Weatherburn, 2002). However with behaviours that can occur at act-level, at partnership-level or at individual-level, it is important to note how this affects the risk behaviour estimate: a study of high-risk HIV-positive men based in San Francisco noted that individual-level data did not show seroadaptive behaviours but

partnership level data demonstrate avoiding anal intercourse, serosorting and strategic positioning being associated with partner serostatus (McConnell, Bragg, Shiboski, *et al.*, 2010).

In a qualitative study (Wolitski, Bailey, Leary, *et al.*, 2003) looking at HIV-positive men from New York and San Francisco, 72% of the 250 men interviewed brought up responsibility for preventing HIV transmission during their interviews with two thirds of them talking about personal responsibility over the partner's health whilst a quarter talked about a shared responsibility and few (22/180) perceived the responsibility to be that of the partners. Interestingly disclosure was seen by some to move the responsibility (for the risk of transmission) to the partner. In a qualitative study from the US (described as medium-sized city in the North East) HIV-positive men attending infectious disease clinic were interviewed on their views on the importance of prevention programmes. The most prominent priorities were for maintaining good mental health, help with substance abuse and advice on stress coping mechanisms, and the participants thought that health promotion programme focusing solely on safer sex would not be welcomed or perceived as effective. Instead the participants hoped for support groups that would focus on life management and be more holistic in its approach (in that context safer sex could be one of the topics). Stigma around HIV-positivity and being gay were also brought up (Vanable, Carey, Brown, *et al.*, 2012). One study also looked at the role of vengeance in HIV-positive men's risk behaviour but based on the results it would seem that revenge is not a significant contributor to unsafe sexual practices in HIV-positive men (Moskowitz & Roloff, 2008).

Barebacking as a behaviour (unprotected anal sex), and barebacking as an identity (men who identify themselves as having intentional unprotected sex) has received growing attention; the concept of intentional unprotected sex among MSM has become more common since the late 1990s (reviewed in Parsons and Bimbi 2007). Berg's review (2009) of barebacking literature comments on defining barebacking as an identity premature, as participants responding to be barebackers does not necessarily mean they associated this with an identity in the same terms as the researchers. In Parsons' study (2007) those HIV-positive MSM who self-identified as barebackers were more likely to report all types of unprotected anal sex (with or without ejaculation, receptive and insertive) than other HIV-positive MSM; they were also more likely to report methamphetamine use and scored higher on substance use during sex as well as knowing more friends who

engage in UAI. In another study HIV-positive barebackers did not report more depressive symptoms than HIV-negative barebackers. HIV-positive barebackers reported more UAI than HIV-negative but depressive symptoms were not associated with the high risk sex measured in the study (Houston, Sandfort, Dolezal, *et al.*, 2012). Based on the literature barebacking was dealt as separate phenomenon from seroadaptive behaviours, and it could be seen as an alternate strategy for the seroadaptive behaviours. Parsons (2007) suggested that to those HIV-positive men who are "consistently barebacking", promoting serosorting or strategic positioning might be a beneficial risk reduction approach.

There is some indication that promoting seroadaptive behaviours might be a feasible strategy among HIV-positive MSM who reported unprotected intercourse with a casual partner based on an intervention in four cities in the United States. The intervention arm received counselling on life quality and coping strategies, sexual health behaviours, such as disclosure and negotiating on safer sex, and on health behaviours, and they reported more serosorting at follow-up than the control arm (Morin, Shade, Steward, *et al.*, 2008). However, another study by Wolitski *et al.* (2005) proposes that promoting seroadaptive behaviours may have some side-effects: they conducted a peer-led intervention among HIV-positive men in New York City and San Francisco, which aimed at reducing unprotected sex with HIV-negative/unknown sex partners and to increase disclosure of HIV status. The intervention arm (with 6 sessions) reported less UAI than the control arm (with one session) at 6-month follow-up but the difference was not a statistically significant. The authors discuss that given the intervention was designed to reduce the risk behaviour among high-risk HIV-positive men, it might have introduced low-risk men to the "ideas" of more risky sex given the intervention was given in a group setting where some participants expressed their beliefs that HIV prevention is the responsibility of the HIV-negative; thus, the authors speculate, the intervention may have had a reverse impact on some of the participants

4.5.5.4 Drug use

Broad literature exists of substance use and increased risk behaviour: for HIV-positive MSM all types of substance use (alcohol, marijuana and the most common "party drugs") in the past 3 months were associated with UIAI and URAI with partners of all serostatus. In a multivariate model methamphetamine and gamma hydroxybutyrate (GHB) were positively associated with IUAI and/or RUAI in all partner serostatus categories.

Barbiturates/tranquilisers and poppers were also associated with IUAI and RUAI with HIV-positive partners (Purcell, Moss, Remien, *et al.*, 2005).

Methamphetamine use has been described in number of studies with the use being consistently shown to increase risk-behaviour and UAI (with HIV-positive men also reporting more methamphetamine use than HIV-negative men) (Pantalone, Bimbi & Parsons, 2008; Chen, Raymond, Grasso, *et al.*, 2012; Forrest, Metsch, LaLota, *et al.*, 2010; Halkitis, Green, Remien, *et al.*, 2005). In HIV-positive substance using MSM consistent serosorting (here defined as concordant partnerships) was associated with high coping self-efficacy (level of confidence of coping under stressful situations) and inversely associated with cognitive escape (use of substances and sex to escape from awareness of norms in sex and HIV) (Kurtz, Buttram, Surratt, *et al.*, 2012).

4.5.5.5 Viral load

Conflicting evidence exist for the use of viral load as a seroadaptive tool. In a meta-analysis of prevalence of UAI in HIV-positive MSM Crepaz *et al.*, (2009) did not find an association between being on ART or having an undetectable viral load and having engaged in UAI. Other identified studies also add to the uncertainty between viral load and decisions to unprotected sex.

When predictors for UAI were measured in serodiscordant male couples, undetectable viral load and ART optimism were found to be positively associated among MSM in Sydney (Van de Ven, Mao, Fogarty, *et al.*, 2005). Also from Australia, a study by Rawstorne et al. (2007) found that HIV optimism and Viagra use were associated with sometimes having UAI with casual partners (when compared to men who reported never doing so). However, in another study undetectable viral load was not associated with unprotected anal sex with the main partner but it was associated with unprotected sex with casual partners (Vanable, Ostrow & McKirnan, 2003). Men who had an undetectable viral load were more likely to discuss this with their sex partners however this had no effect on having potentially sdUAI partners (Guzman, Buchbinder, Mansergh, *et al.*, 2006).

Table 4. Summary of intrapersonal factors contributing to seroadaptive behaviours in HIV-positive MSM, and their hypothesised association with transmission risk.

Intrapersonal factors	Potential relation to seroadaptive behaviours (SABs) in HIV-positive MSM	Hypothetical relation to HIV/LGV transmission risk
Time since diagnosis	SABs may be time-dependent.	Changing transmission risk profile over time.
Intention and self-efficacy	Likely to be an important determinant for consistent SABs and practicing safer sex.	Likely to reduce transmission risk for both HIV and LGV.
Risk behaviour	Related to number of other factors, likely to make consistent SABs less likely.	Good evidence for increased risk for both HIV and LGV.
Drug use	Increased overall risk behaviour.	Good evidence for increased risk for both HIV and LGV..
Viral load	It is not clear how widely viral load is used as a SAB.	For those with regular viral load monitoring, could reduce HIV transmission; but might increase LGV risk if sex is unprotected.

4.5.6 Proposing a conceptual framework

Based on the literature review, I propose a conceptual framework, which is presented in Figure 2. The framework is divided into structural, community, interpersonal and intrapersonal factors that have been hypothesised to be linked to seroadaptive behaviours in the literature. In the literature a varying degree of information was available on each of the factors and their level of influence, but the framework illustrates how seemingly distal factors, such as geographical region or the venues in the community can have an effect on seroadaptive behaviours. The factors can have a direct effect on seroadaptive behaviours (POZ parties, searching for HIV-positive sex partners online, self-efficacy to negotiate for safer sex) or in an indirect manner (drug use which interferes with intentions, sex-on-premises venues where non-disclosure is a norm).

The framework allows us to investigate hypothetical ways in which seroadaptive behaviours have aided LGV re-emergence. Even when seroadaptive behaviours are “successfully” practiced (reduced risk of HIV transmission) LGV transmission risk may still remain. There are also a number of situations where overall risk

of STI and HIV transmission might be increased (drug use, increased number of partners, sex-on-premises venues, HIV-related stigma).

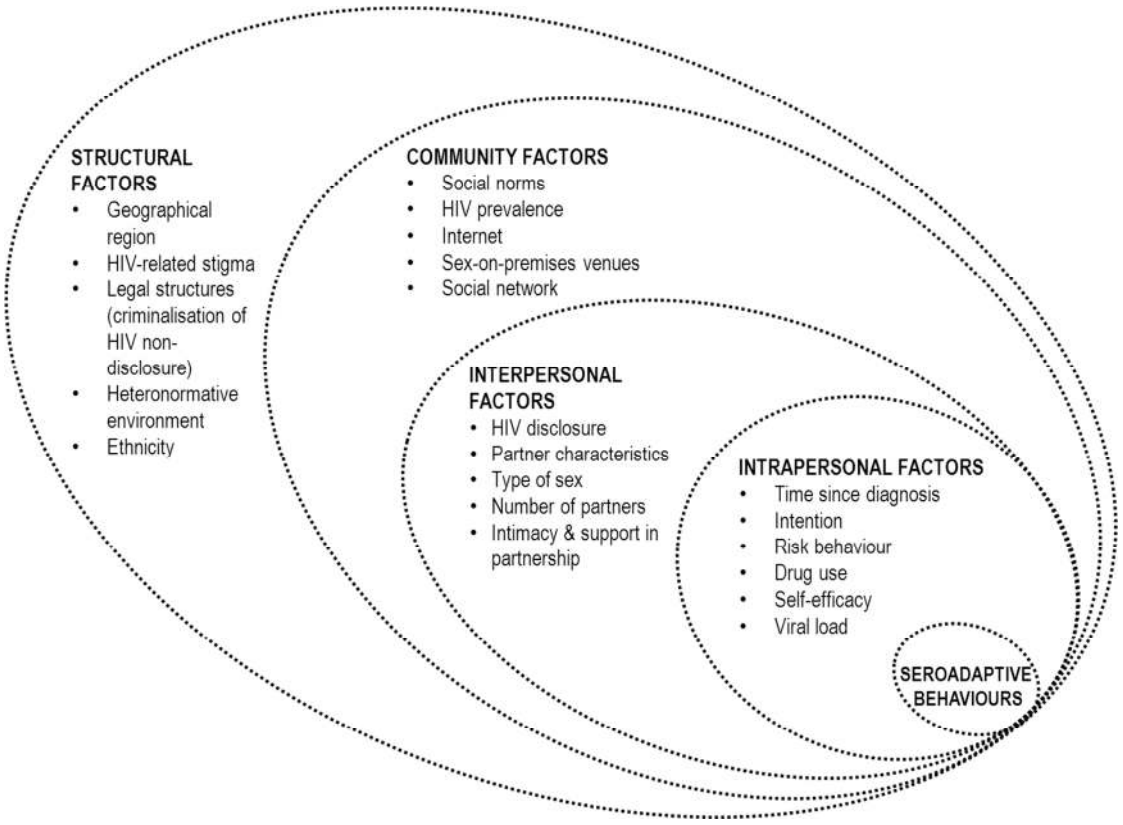


Figure 2. Conceptual framework of the social epidemiology surrounding seroadaptive behaviours in HIV-positive MSM.

4.6 Discussion

This chapter narratively synthesises the literature on seroadaptive behaviours in HIV-positive MSM, and suggests ways in which seroadaptive behaviours can be hypothesised to be linked to HIV and LGV transmission risk. Based on the review a conceptual framework was proposed.

Seroadaptive behaviours, at least in academic research, are a phenomenon limited to high-income countries, presumably partially driven by the more visible and recognised communities of MSM, good access to HIV testing, higher (diagnosed) HIV prevalence, awareness of HIV risk factors. This is supported by the few Asian-based studies where the prevalence of HIV was low in the populations sampled, knowledge of HIV among MSM low and serodiscordant relationships common.

A theme which emerged in the literature was that seroadaptive behaviours are being employed on partnership level with regular partners. When the anonymity of a sexual encounter increases (or when the number of partners increases), the perceived responsibility to disclose and discuss sex wanes, and in some instances the responsibility for safer sex is shifted to the partner. A prime example of this are the most high-risk settings (for sdUAI and for HIV prevalence), such as bathhouses and sex parties, where disclosure is seen as nonessential and where a separate "micro culture" seems to prevail.

Sheon and Crosby's study from San Francisco (2004) described the dominance of HIV-positive men where the HIV-related stigma was absent, whilst in other settings with lower HIV prevalence HIV-positive men felt disempowered. This "sexual hierarchy" would support Factor's resistance framework where a minority status is a component of HIV-related stigma and resulting outcomes, and possibly a contributor to risky sex as a form of defiance against majority culture values. This is also suggested in the discussion of a Sigma Research report of high-activity men; the authors speculate the transgressive aspect of homosexual sex is thrilling in itself and that leads some men to seek more extreme forms of sex, and for them the risk and taboo aspects of UAI will increase its attractiveness (Davies, Reid & Weatherburn, 2002).

The aetiology of seroadaptive behaviours is not obvious based on studies conducted thus far and the cross-sectional nature of the majority of studies further confounds causal inference. The strong assortative mixing between HIV-positive MSM can potentially be a conscientious risk management strategy, but as Murphy *et al.* (2012) discuss and as Chaudoir's HIV Disclosure Process Model (Chaudoir, Fisher & Simoni, 2011) indicates, social support and shared experiences gained from an HIV-positive partner can be a significant factor in driving HIV-positive men into a "seroadaptive environment" rather than avoiding transmission of HIV as the primary component in directing behaviour.

It is important to note that although the majority of the review articles specifically focused on high-risk MSM it was clear that a proportion of HIV-positive men abstain or limit their sexual activity, whilst others have a consistent safe sex approach or they are in a monogamous partnership.

It was also evident that seroadaptive behaviours lack standard definitions. We can hypothesise that seroadaptive men may be broadly divided into – similar to the serosorting definition described in (Marcus, Schmidt & Hamouda, 2011) – strategic 'seroadapters' (preferred method and part of partner selection approach) whilst some are tactical seroadapters (where the behaviour is context specific and maybe part of a more complicated sexual decision-making process). Many men may also be "accidental" seroadapters whereby the context of the sexual encounter is seroadaptive rather than the partnership-level behaviour. Too strict definitions of what constitutes seroadaptive behaviours would be impractical and fail to understand the variety and contextual nature of behaviour. However a too broad definition – such as the definition presented in the introduction where any modification of sexual behaviour based on one's own, their partner's or the perceived HIV transmission risk is seroadaptive (Le Talec & Jablonsky, 2008) – is difficult to translate into epidemiologically meaningful concept or a public health message.

It is therefore difficult to ascertain that seroadaptive behaviours, as a whole, lower the risk for HIV transmission, but it would seem that any behavioural strategy is better than not having any at all and the most worrying group from health protection point of view would be men who are not able to manage their sexual-risk taking and lack the self-efficacy to plan and maintain some form of safer sex approach (causing harm to themselves and to others).

From the qualitative work among HIV-positive men, some expressed a need for programmes that would focus on life management and not solely on sexual health. Also, given the diversity of experiences among HIV-positive men, considering the dynamics of the specific community is important to be able to adjust the message to the prevailing norm. We should aim to understand better how HIV-positive MSM understand seroadaptive behaviours, how they employ them, and whether some HIV-positive MSM (such as those who report frequent UAI) would be able to reduce their risk (for HIV transmission and STI acquisition) by adopting seroadaptive strategies. Given the strong involvement of community level norms and venues that can maintain the seroadaptive behaviours, an ideal intervention would be community-based (given the number of factors affecting individual-level determinants, it is easy to see how interventions targeting the behaviour of an

individual are challenging). In the beginning of the HIV epidemic, gay communities approached HIV as a disease that affected the community. The intervention could, for example, utilise Factor's resistance framework, and the unity of gay community in promoting a shared responsibility over HIV and STI transmission (Factor, Kawachi & Williams, 2011). In practice this would be difficult to implement in a metropolitan setting such as London where the diversity of MSM communities is greater.

From the perspective of LGV re-emergence in the UK, seroadaptive behaviours form a context which can facilitate the establishment of an STI. Even though majority of the seroadaptive literature comes from the United States, the experiences described by the HIV-positive men in Relative safety II (Bourne, Dodds, Keogh, *et al.*, 2009) were very similar to those described elsewhere in the literature. Locations such as Manchester and London may be central to HIV-positive subcultures. Serosorting was widely reported with authors noting that seroguessing seemed to approximate serosorting in many instances, and using online websites was a common method to mitigate the situation. Different tactics were described for romantic partners compared to casual partners with disclosure seen as important if the relationship was going somewhere. Thus the social epidemiological framework of seroadaptive behaviours developed here could be applicable to the UK situation.

4.6.1 Strengths and limitations

Given the lack of precise definitions and the broad subject areas reviewed, the literature search was not exhaustive, and this study cannot be called a systematic review. The study is limited by its broad goals, and the literature review was able to scrape the surface of the complex phenomena it describes. The review could have been made more systematic by looking at a specific seroadaptive behaviour in detail, but this review's strength is that it looks at the variety of behaviours, as they rarely occur in isolation, and it is thus able to capture some of the problems in the studies, and the diversity of experiences.

To systematically cover all the levels several literature reviews should have been incorporated with level-specific search words. For example, there is indication for a connection between seroadaptive behaviours and HIV prosecutions, but a broader literature exists on the topic where coping strategies for the legal structures are discussed but seroadaptive behaviour terminology is not used. Also the minority status and Factor's

resistance framework are interpreted to be underlying factors in seroadaptive behaviours, but studies have not specifically focused on this. Therefore the conceptual framework, although based on published studies, is not validated, so its main function is to be a platform for further hypotheses. Narrative reviews are subjective in nature, and can be used to say many things of the existing literature. Acknowledging this, the approach used here focused on the experience of HIV-positive MSM trying to understand reasons and situations where seroadaptive behaviours occur.

In research synthesis, all biases present in individual observational studies, remain in the meta- analysis (Egger, Davey Smith & Altman, 2001). No quantitative analysis was employed here, and heterogeneities between studies were not explored. It is also important to note that the majority of studies required their participants to be sexually active, and many focused on high-risk setting. Thus those men who consistently practice safe strategies or who are not sexually active are most certainly underrepresented in this review.

4.7 Conclusions

This review showed a pattern of seroadaptive behaviours which facilitate relatively assortative sexual mixing by HIV-positive men. Given the high levels of UAI, this can explain how STIs are seeded into HIV-positive sexually active population. Assortative mixing between HIV-positives is likely to decrease the onward transmission of HIV, which was hypothesised as reason for the increased STI diagnoses while HIV incidence had stabilised in San Francisco (Truong, Truong, Kellogg, *et al.*, 2006). However many HIV-positive men, especially in sex-on-premises venues, report sdUAI, and this review alone is not enough to explain why LGV is not more common in HIV-negative high-risk men.

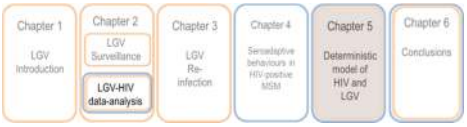
Further studies looking at seroadaptive behaviours should better assess intention, and whether and how seroadaptive behaviours are employed as a strategy. This would enable suitable prevention programmes for MSM who have seroadaptive behaviours. The biggest intervention challenge is still to countervail forces of peer norms and drug abuse in these small sub-communities of MSM to whom traditional prevention messages are ineffective.

This review helps us to better understand the contextual evidence available for why and how seroadaptive behaviours occur and how they may have contributed to the re-emergence of STIs, such as LGV. It also gives a more multidimensional perspective to the sexual lives of HIV-positive MSM instead of the rather superficial description of high-risk behaviour that is more often attributed as the cause of STI transmission.

Chapter 5

Serosorting, high-risk behaviour and the intersecting epidemics of HIV and LGV:

A conceptual mathematical modelling approach



5.1 Summary

The previous chapters have used statistical analysis and social epidemiological methods to understand the association between HIV and LGV. I conducted a review on seroadaptive behaviours in HIV-positive MSM, which showed that HIV-status affects how individuals select their sexual partners. This chapter will further explore this phenomenon by aiming to bridge the gap between the observed seroadaptive behaviours and hypothesised link between serosorting and LGV-emergence. This is done by adopting a mathematical modelling approach to explore the dynamic transmission processes of LGV and HIV epidemics focusing on interaction due to behaviour.

I present a deterministic model incorporating HIV and LGV which includes mixing by perceived HIV status (serosorting) and sexual activity class. This was used for theoretical examination of how serosorting and activity mixing were associated with varying epidemic sizes for HIV and LGV.

I will demonstrate that HIV and LGV infection will tend to concentrate in the same population irrespective of epidemic size or mixing patterns due to their same mode of transmission. This association is further strengthened in a low level LGV epidemic and with a high-level of serosorting. Serosorting is also affected by the amount of heterogeneity in the population sexual activity.

The mathematical model in this chapter demonstrates that there is always going to be some overlap between infections with the same mode of transmission, but it gives further support to the hypothesis that serosorting can explain the particularly high HIV prevalence in those who acquire LGV.

5.2 Introduction

The purpose of this chapter is to explore the hypothesis that serosorting has influenced the emergence of STIs like LGV. I am using a mathematical model incorporating HIV and LGV transmission in an MSM population where sexual mixing is based on sexual activity class and perceived HIV status. In the previous chapters I have explored the association between HIV and LGV through statistical analysis of LGV Enhanced Surveillance data. The analyses demonstrated a diagnostic bias in favour of the HIV-positive MSM but also that HIV-positive men were more likely to report unprotected receptive anal sex before their LGV episode when compared to HIV-negative/unknown LGV cases. In a second analysis of repeat LGV episodes HIV-positive men were more likely to be diagnosed with a repeat infection. Repeaters were more likely to present in their first LGV episode with a concurrent hepatitis C infection and gonorrhoea when compared to those who did not have a reported re-infection.

It has been suggested that serosorting and sex-on-premises venues could be the reason for the formation of dense sexual networks which has facilitated the re-emergence and persistence of LGV (Ward, Martin, Macdonald, *et al.*, 2007; van de Laar, 2006). To better understand the social context I looked at seroadaptive behaviours in HIV-positive MSM through a literature review and proposed a social epidemiological framework. Reporting serosorting has been associated with self-reported STIs (Marcus, Schmidt & Hamouda, 2011), and the literature demonstrated that seroadaptive behaviours are adopted by MSM in high-income countries. However, there is no direct evidence in the LGV Enhanced Surveillance to suggest the HIV-positive LGV cases preferentially seek out HIV-positive sex partners. Therefore I am using a mathematical model to better understand how serosorting might influence and support LGV re-emergence. In the model LGV and HIV do not interact biologically in relation to transmission, and the association between the two is based on behaviour.

Mathematical models are a simplification of reality and they are able to provide a laboratory-like environment where we are able to define and control every aspect of the population and the infection(s) we are interested in. This allows us to investigate the dynamic, and often non-linear, processes that affect infectious diseases simultanously, such as demographic change, behavioural patterns in the population, natural history of the

disease and infection transmission (as a function of the prevalence of the infectious individuals, risk of infection and availability of susceptible individuals.). In the face of complexity, mathematical models tend to be reduced to including essential features of the system in light of the research questions being addressed. Therefore the mathematical modeller has to come to a compromise in desire for accuracy (complexity of the model), transparency (ability to understand how the model functions and how each component affects the dynamics) and flexibility (adaptability of the model to new situations) (Keeling & Rohani, 2008; Garnett, Cousens, Hallett, *et al.*, 2011).

Mathematical models have been used to explore social and behavioural aspects of infectious diseases as reviewed by Cassels & Goodreau (2011) and they have also been used to examine structural factors such as how sex workers' role as a core group is dependent on their duration in sex work and the potential role of the men who control the sex industry as a reservoir of infection (Watts, Zimmerman, Foss, *et al.*, 2010) and health care provision such as in the mathematical model looking at the impact of demand and supply for GUM clinic services on the incidence of gonorrhoea (White, Ward, Cassell, *et al.*, 2005). There are several models that have focused on chlamydia transmission in heterosexual population (Vickerman, Ndowa, O'Farrell, *et al.*, 2009; Kretzschmar, Turner, Barton, *et al.*, 2009), and mathematical models have also incorporated co-infections such as HSV-2 and HIV (Foss, Vickerman, Chalabi, *et al.*, 2009).

Most mathematical models of HIV in MSM have focused on estimating HIV projections using behavioural or biological intervention (reviewed by Punyacharoensin, Edmunds, De Angelis, *et al.*, 2011). The review showed that complexity of model structure has increased over time and mathematical models for HIV in MSM are increasingly incorporating sexual activity, age and HIV treatment related characteristics. Mathematical models have been able to demonstrate the difference in HIV transmission potential between MSM and heterosexual populations, and the importance of higher transmission probability in receptive anal sex (compared to vaginal intercourse) and the role versatility in MSM (compared to role segregation in heterosexuals) (Goodreau & Golden, 2007).

Mathematical models have also been employed to examine the impact of social phenomena on disease transmission, such as the role of commercial sex venues for HIV transmission in MSM in Seattle (Reidy & Goodreau, 2010). The study showed how closing venues would be unlikely to reduce HIV incidence if sexual partnerships are sought from elsewhere (shifting the sexual mixing patterns towards more random mixing). Mathematical models have also been used to investigate serosorting as a primary HIV prevention strategy (Golden, Stekler, Hughes, *et al.*, 2008; Wilson, Regan, Heymer, *et al.*, 2010). The work by Golden *et al.* estimated serosorting to be a protective of HIV in their population of MSM in Seattle whilst the study by Wilson *et al.* found that for serosorting to be an effective primary prevention strategy the proportion of undiagnosed MSM needs to be below 20%. In this work, I will examine how serosorting – along with other behavioural and biological parameters – affect HIV- and LGV prevalence.

Compartmental models assume implicitly that partner change rate is a major determinant in transmission. Partner change rate therefore becomes both the direct and indirect measurement of sexual risk behaviour and its underlying determinants (for example drug use which may disinhibit one's sexual conduct or other social and structural features that may increase a person's sexual risk-taking or vulnerability to STI acquisition). This simplification has been criticised for making assumptions of the high-activity group which may or may not be based on empirical evidence (Cassels & Goodreau, 2011), and it is a structural limitation of compartmental models. Further discussion on this limitation in relation to this model is in the methods (section 5.4.3).

As the previous chapter demonstrates, sexual mixing between HIV-positive men is a complex phenomenon and highly dependent on the social setting and community norms: there is evidence that HIV-positive and HIV–negative men meet partners through different venues (Grov, Golub & Parsons, 2010), and some venues have been created or have transformed to cater for the HIV-positive and/or high-risk MSM population. In a study by Velter et al. (2009) serosorters reported more sauna, backroom and cruising ground venue attendance than non-serosorters (however the study was not stratified by HIV status of the respondent). Furthermore, HIV prevalence differs according to venue type with a cross-sectional study in California circuit parties having 40.8% HIV prevalence compared to 20.0% in gay bars and clubs (Xia, Tholandi, Osmond, *et al.*, 2006). In POZ parties for HIV-positive men, the self-reported HIV prevalence was 95% (Clatts, Goldsamt & Yi, 2005). In

comparison HIV prevalence in MSM in London was estimated as 12.3% in a venue sample (Dodds, Johnson, Parry, *et al.*, 2007), PHE estimates it to be 9% in London (Delpech, 2012) and Sigma Research report for London in 2008 had 15.4% of its respondents reporting having tested HIV-positive (Sigma Research, 2009).

The correlation between seroadaptive behaviours in HIV-positive MSM and increase in the STIs in this group seem sensible on a population level, but to better understand the interactions between the two infections and the extent to which behaviour, and particularly serosorting, could explain the association between LGV re-emergence an underlying model of transmission is needed. In the UK, the LGV epidemic curve has shown low case numbers that have fluctuated over time. This stochastic trend of LGV in the UK data means that it would be difficult to reproduce the epidemic; even with identical conditions, random variation would lead to quantitatively different LGV patterns. In addition, there is limited understanding of the natural history of LGV or the necessary biological parameters to construct a reliable LGV model. The aim of this modelling is not to reproduce the observed epidemic but to understand how behaviour and sexual mixing-patterns contribute to the association between LGV and HIV.

5.3 Objectives

The overall aim of this chapter was to investigate how the epidemics of HIV and LGV interact using a mathematical model, and more specifically to explore the possible role of serosorting as an explanation for the high-level of LGV-HIV co-infection.

Two specific objectives were defined to achieve this:

- To assess how the underlying HIV prevalence in the population affects HIV prevalence in LGV cases.
- To estimate the extent to which sexual mixing by serostatus (serosorting) and by activity class affect HIV prevalence in LGV cases in the absence of biological synergy.

5.4 Methods

5.4.1 Model structure and natural history of infection

A flowchart for the model is presented in Figure 1 with the HIV progression model presented in blue. A list of parameter estimates, including references they are based on, is shown in Table 1. HIV model structure is designed from a behavioural, rather than clinical point of view. There are four stages (assigned letter *I*): HIV-negative, HIV infected who are unaware of their infection (perceived HIV-negative), HIV-positive who are aware of their infection (diagnosed HIV-positive), and HIV-positive receiving ART. In the model simulations each of the HIV-infected stages can have a different transmission probability which approximates the different clinical stages. HIV-infected unaware of their infection are diagnosed at rate ω_1 and HIV-diagnosed initiate ART at a rate ω_2.

This HIV-model is then coupled with an LGV model (presented in orange in the figure) with an SIS (susceptible-infected-susceptible) structure where individuals in each of the HIV-model compartments can acquire LGV, and become infectious; following recovery these individuals return to susceptible category and can become re-infected[27]. HIV-negative LGV-infected individuals can also acquire HIV-infection during their LGV infection, and they can proceed through the HIV compartments whilst LGV-infected. Those in the HIV-negative LGV susceptible compartment can acquire HIV, LGV or both at the same time.

In the results of the data-analysis presented in chapter 2, I found evidence of diagnostic bias depending on HIV status; therefore there is a different recovery rate of LGV for HIV-negative (σ') and HIV-diagnosed (σ). For the HIV-positive undiagnosed compartment one can either get diagnosed with LGV and HIV at the same time (which also results in treatment of LGV, σ'), or a smaller rate (σ´´) is assigned to take into account the small

[27] As discussed in the first chapter, the natural history of LGV is poorly understood. In the vast majority of LGV cases LGV has been highly symptomatic. It is not clear whether some LGV remains asymptomatic or if time to diagnosis – as hypothesised in a review by White (2009) – is more influential for the development of symptoms. Given the lack of previous LGV models, and re-infections that are observed in the Enhanced Surveillance, I used an SIS model to approximate LGV natural history.

proportion of LGV-infected who would refuse HIV-testing and get treated with LGV without HIV-diagnosis. The recovery rate in the model is assumed to be driven by treatment seeking due to LGV-related symptoms.

The model population is stratified into three activity classes (*i*), which are differentiated by having different partner change rates: high, medium and low-activity individuals. Partner change rate remains the same regardless of infection status, and does not vary post HIV-diagnosis, but the mixing matrix is different for those who are HIV-negative/unaware of their infection (*l*=1 or *l*=2) than for those who know they are HIV-positive (*l*=3 or *l*=4). There is entry (*b*) and exit (μ) from the model describing the duration of sexual activity, which is roughly estimated by the age range of LGV infected in the surveillance data, with everyone entering the model as HIV and LGV uninfected (thus external sources of infection are not considered), and primary source of entry and exit are assumed to be age-related. The exit rate is equal for all infection stages, and the population size in the model is kept constant.

In this study LGV and HIV do not interact biologically in relation to transmission. LGV infection is not assumed to directly affect the progression of HIV-infection; however undiagnosed HIV-infection can be diagnosed due to individuals seeking care for LGV infection, moving the individuals from *l*=2 to *l*=3. This makes diagnosing HIV faster after LGV is introduced which affect the proportion who are unaware of their infection and the overall force of infection for HIV[28]. During the model simulations the effect of this phenomenom was monitored.

[28] Leaving this out would keep LGV and HIV completely separate in terms of their transmission dynamics, but it would also make the model more unrealistic if those with undiagnosed HIV are not diagnosed during LGV treatment.

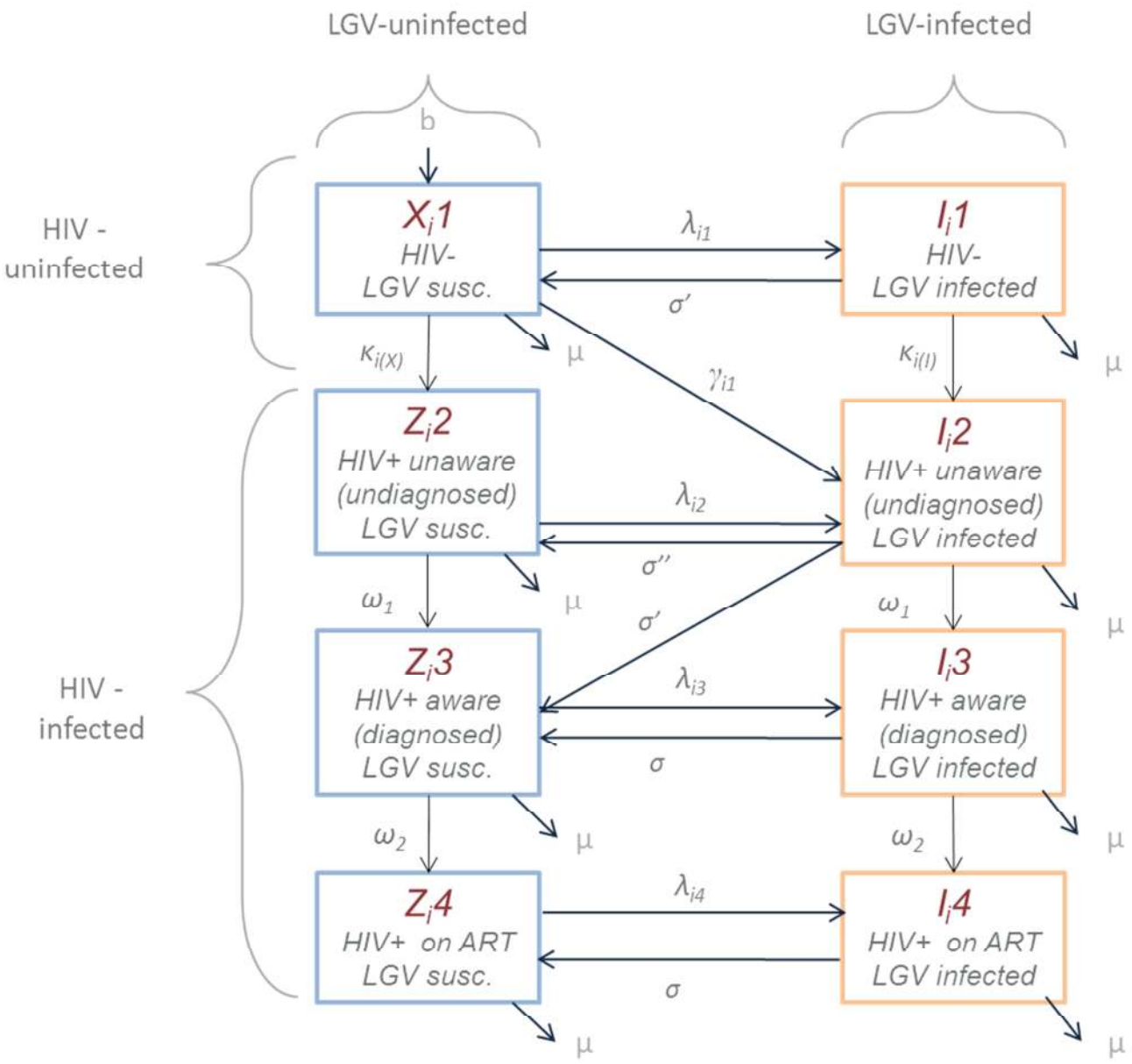


Figure 1. Flowchart of HIV-LGV model

Table 1. Table of parameters, their time dependency and the distribution of the parameter range (along with minimum and maximum) for those parameters that are used for Latin Hypercube sampling. For the parameters that are fixed in all model simulations, the parameter estimate is given.

Parameter / variable name	Description	Time dependency	Distribution	Estimate	Min	Max	Source / explanation
N	MSM population size in London	Constant	Fixed	70000			Approximating the size of MSM population in London, see footnote 29
b	Inflow/ month	Constant	Uniform		1/294 (~40yr)	1/588 (~20yr)	Wide age range in LGV Enhanced Surveillance of 18-67, time in the model varied between 20 and 40 years.
	Outflow = inflow						Constant population size
$X1_i$	Fraction HIV-	Adjusted (t)	Fixed	0.95			Initial conditions for the model: 95% of population is HIV-susceptible
$Z2_i$	Fraction HIV+ unaware	Adjusted (t)	Fixed	0.001			Initial condition for the model: HIV is seeded at 5% prevalence into the population
$Z3_i$	Fraction HIV+ diagnosed	Adjusted (t)	Fixed	0.02			
$Z4_i$	Fraction HIV+ on ART	Adjusted (t)	Fixed	0.029			
	Fraction high-activity	Constant	Uniform		0.02	0.12	Sexual activity parameters are estimated based on Sigma Research, London 2008, (Sigma Research, 2009). See Table S3 in appendix.
	Fraction medium-activity	Constant	Uniform		0.121	0.2	
	Fraction low-activity	Constant	Uniform	1-high and medium-activity			
c1	pcr for high-activity /month	Constant	Uniform		0.42	2	Number of male UAI partners last year (see Table S3 in appendix) (5-24UAI partners/year)
c2	pcr for medium-activity /month	Constant	Uniform		0.22	0.41	(2.6-4.9 UAI partners per year)
c3	pcr for low-activity /month	Constant	Uniform		0.03	0.21	(0.4 -2.5 UAI partners per year) (Sigma Research, 2009)
θ	pcr compromise, assumed equal	Constant	Fixed	0.5			Adjusted according to (Garnett & Anderson, 1994)
ε_1	Mixing parameter for perceived HIV status of partner	Constant	Uniform		0	1	Varied from fully assortative (0) to proportionate mixing (1)
ε_2	Mixing parameter for activity class	Constant	Uniform		0	1	Varied from fully assortative (0) to proportionate mixing (1)

β_{HIV1}	Transmission probability per partnership, dependent on the infection status of the HIV+ (unaware HIV+)	Constant	Uniform		0.01	0.158	Adjusted summary estimates for transmission probability per partnership in URAI (10.0%; 95% CI 4.2, **15.8**) and IUAI (0.7%; 95% CI **0.0**, 1.3) from a meta-analysis (Baggaley, White & Boily, 2010); minimum and maximum estimates are adapted from the extremes of the 95% CIs
β_{HIV2}	Transmission probability for aware/diagnosed HIV+	Constant	Uniform		0.01	0.158	''
β_{HIV3}	Transmission probability for HIV+ on ART	Constant	Uniform		0.01	0.22	Modelling paper by (Hallett, Smit, Garnett, *et al.*, 2011): Varies from 1-22% depending on condom use in long term partnerships; In MSM on HAART, HIV was detected in 30% of semen samples (associated with STI urethritis and UAI with HIV+ partner) (Politch, 2012)
ω_1	Rate of HIV diagnosis/ month	Constant	Uniform		1/100.8 (8.4 yr)	1/31.2 (2.6 yr)	Birrell et al., (2013) estimate that the mean time to diagnosis has declined from 4 years (95% CrI 3.8, 4.2) in 2001 to 3.2 years (95% CrI 2.6, 3.8) by end of 2010. Minimum was selected from the minimum estimate in the credibility interval (CrI) range (2.6) and the maximum was double their maximum estimate in 2001 (4.2) to account for those who have never tested.
ω_2	Rate of initiation with ART / month	Adjusted (t)					To be set in a way that the proportion on ART remains stable: 70.7% of HIV+ MSM in North East London are on ART (Elford, Ibrahim, Bukutu, *et al.*, 2007); in SOPHID approximately 70% of MSM are on ART (of those with information available) (SOPHID)
LGV specific parameters							
β_{LGV}	Transmission probability per partnership for LGV	Constant	Uniform		0.2	0.8	An assumption Quinn et al., (1996) estimate that transmission probability for chlamydia per heterosexual partnership is 0.68), and Holmes *et al*. STD book considers LGV to be less transmissible than gonorrhoea (Stamm, 2008b).
φ	STI cofactor for HIV transmission	Constant	Fixed	1			
α	HIV cofactor for LGV transmission	Constant	Fixed	1			
σ	Recovery rate for known HIV+/ month	Constant	Uniform		1/3	1/5.88	An assumption From 3 to 6 months; I assume everyone will eventually become symptomatic and seek treatment. In LGV ES data duration of symptoms has a mode of one week with HIV-negative reporting a longer duration of symptoms than HIV-diagnosed. See section 5.4.11 for further detail. For comparison, a mathematical modelling study estimated untreated symptomatic *C. trachomatis* recovery time to be 17 weeks (95% CI 4.5, 35.1 weeks) (Johnson, Alkema & Dorrington, 2010)
σ'	Recovery rate for HIV-/month	Constant	Uniform		1/3	1/5.88	
σ''	Recovery rate for undiagnosed HIV+ back to unaware / month	Constant	Uniform		1/3	1/5.88	
LGV seed	Number of initial infections		Fixed	10			

5.4.2 Parameters

In this model I am first introducing HIV into the population and the model simulations are repeated for the parameters varied, described in Table 1, by their distribution (fixed or uniform) and the estimate (for fixed parameters) or the minimum and maximum estimates for the uniform distribution. I am using Latin Hypercube Sampling (section 5.4.9) which generates a range of model simulations, which vary in terms of parameter estimates used. After HIV has reached equilibrium prevalence (discussion of this method in 5.4.10) LGV is introduced to the model using a second Latin Hypercube Sampling for the LGV natural history parameters. This means that the behavioural parameters are selected based on the HIV simulations.

The opportunity for an epidemic to occur (for $R_0>1$) is dependent on the transmission probability, duration of infection, and partner change rates in the population (and these parameters are therefore correlated). Given the limited understanding of the natural history of LGV I assume the duration of infection is not very long, given most LGV becomes increasingly symptomatic over time making it likely for patients to seek care – and it does not seem to be highly transmissible infection either (based on its low prevalence estimates overall (Ward, Alexander, Carder, *et al.*, 2009; Annan, Sullivan, Nori, *et al.*, 2009) despite the levels of partner change rate reported by LGV cases. However there is a great deal of uncertainty around the natural history parameters for LGV. Therefore, I will allow for variation for the natural history parameters for LGV but I am focusing on the partner change rate and proportion of high-activity individuals in the population.

In this study most parameters are varied using uniform distribution with Latin Hypercube Sampling. This is done to reflect the exploratory nature of the study and lack of information on parameter values. The population behaviour approximates the behaviour of MSM in a large metropolitan city, such as London. Estimating the distribution of MSM is challenging[29], but I will look at HIV in equilibrium state and outputs are

[29]Natsal 2000 gives an estimate of 5.5% (95% CI 4.2, 7.2) of men reporting homosexual partners in last 5 years or 1.4% (95% CI 1.0, 2.0) of men who reported 1 or more new homosexual partner in the last year (Johnson, Mercer, Erens, *et al.*, 2001). Using Office for National Statistics (ONS) population estimate from 2003 Greater London had 4,900,900 habitants of working age (Office for National Statistics, n.d.) of which approximately half are males. Using 1.4%-5.5% range to estimate the size of London MSM population we get a wide estimate of 34,300-134,800 MSM in London.

given as a proportion of the population, therefore the model population size is less relevant and a somewhat arbitrary number of 70,000 is selected for the model to approximately represent the MSM population in London.

5.4.3 Considerations for sexual risk behaviour representation in the model

Using the number of sex partners as an indicator of risk of infection is based on the assumption that it is a strong determinant of acquisition risk. However the threshold for a high-activity class is arbitrary, and the reliability of data regarding past sex partners is often subject to biases in recall and reporting.

It is important to distinguish number of sex partners, number of anal sex partners and number of unprotected anal sex partners – and how these affect the interpretation of estimates in relation to a person's risk – since there are sexually active MSM who do not have anal sex, or have anal sex in only certain situations or with certain partners. This was also highlighted by Sigma research in a report of MSM with high number of partners (Davies, Reid & Weatherburn, 2002). However, the report demonstrates that those who reported a high number of partners (defined as more than 30 sex partners in the past year) were more likely to have anal intercourse with both regular and casual partners and more likely to have unprotected anal intercourse with a casual partner as well as report more UAI partners and more likely to report discordant UAI than those who reported less partners; 28.6% of them reported an STI in the previous year. Thus, despite its problems, the number of sex partners reported can serves as a direct and indirect estimate of someone's risk for STIs. In the appendix for chapter 5, I present distributions of sex partners reported through Natsal, Sigma research, and Enhanced Surveillance systems for gonorrhoea, syphilis and LGV. The problem (for parameterising models) with the commonly reported figures is that they often report the total number of sex partners with no information of how many of these were unprotected (see Table S1 and TableS2 in the appendix). For this model I have used the number of UAI partners[30] reported by MSM in London for the 2008 Sigma research as guidance (see Table S3 in the appendix). The estimate is however capped with 7.9% reporting more than 5

[30] Condom use is not incorporated in the model and partner change rate is based solely on UAI. This does not consider other modes of transmission, or role positioning (assuming all men are versatile with an equal likelihood of transmission per partnership).

partners in the past 12 years, which is unfortunate as it is the proportion and numbers reported by those with most partners that drives the epidemic in the deterministic model. For comparison there were 15.4% reporting over 30 sex partners in Sigma 2008, and 7.5% reporting 15-100 partners in Natsal 2000 (both of these are total number of sex partners within the last 12 months). In LGV Enhanced Surveillance 6% reported 30-213 sex partners and in GRASP 8.6% reported more than 11 sex partners in the past 3 months (see Table S2).

5.4.4 Ordinary differential equations

The model is based on ordinary differential equations (ODEs), which are solved through numerical integration. The model was originally coded in Berkeley Madonna using Runge-Kutta4 solver, and then the model was moved to Matlab[31] where inbuilt function exists for Runge-Kutta4 solver (ode45) which was used for numerical integration of the equations in this study. Matlab versions R2010a and R2012a have been used for the mathematical modelling.

HIV-progression is defined by these four sets of ODEs for each three activity groups:

(1)

$$\frac{dX_{i1}}{dt} = N_{il}b + I_{i1}\sigma' - X_{i1}\kappa_{i(X)} - X_{i1}\gamma_i - X_{i1}\lambda_{i1} - X_{i1}\mu$$

(2)

$$\frac{dZ_{i2}}{dt} = X_{i1}\kappa_{i(X)} + I_{i2}\sigma'' - Z_{i2}\lambda_{i2} - Z_{i2}\omega_1 - Z_{i2}\mu$$

(3)

$$\frac{dZ_{i3}}{dt} = Z_{i2}\omega_1 + I_{i2}\sigma' + I_{i3}\sigma - Z_{i3}\lambda_{i3} - Z_{i3}\omega_2 - Z_{i3}\mu$$

(4)

$$\frac{dZ_{i4}}{dt} = Z_{i3}\omega_2 + I_{i4}\sigma - Z_{i4}\lambda_{i4} - Z_{i4}\mu$$

X_{i1} is HIV-negative and susceptible for LGV and Z_{i2}- Z_{i4} are the HIV-infected stages. Parameter b describes the inflow of new individuals to the population and μ is the outflow of individuals, and they are set so that the

[31] When the model was moved from Berkeley Madonna to Matlab the outputs from the two software were checked against each other for accuracy of re-coding.

population size remains constant. κ_i is the force of infection for HIV, γ_i is the force of infection for HIV-LGV simultaneous co-transmission and λ_{il} is the force of infection for LGV defined by HIV-status *l* and activity class *i.* σ is the recovery rate from each LGV infection compartment. ω_1 and ω_2 describe the rate of being diagnosed with HIV and the rate of beginning ART, respectively.

LGV is incorporated into the model through these four sets of equations (for each activity group) that pair with the HIV compartments:

(5)

$$\frac{dI_{i1}}{dt} = X_{i1}\lambda_{i1} - I_{i1}\sigma' - I_{i1}\kappa_{i(l)} - I_{i1}\mu$$

(6)

$$\frac{dI_{i2}}{dt} = Z_{i2}\lambda_{i2} + X_{i1}\gamma_i + I_{i1}\kappa_{i(l)} - I_{i2}\sigma' - I_{i2}\sigma'' - I_{i2}\omega_1 - I_{i2}\mu$$

(7)

$$\frac{dI_{i3}}{dt} = Z_{i3}\lambda_{i3} + I_{i2}\omega_1 - I_{i3}\sigma - I_{i3}\omega_2 - I_{i3}\mu$$

(8)

$$\frac{dI_{i4}}{dt} = Z_{i4}\lambda_{i4} + I_{i3}\omega_2 - I_{i4}\sigma - I_{i4}\mu$$

5.4.5 Force of infection for LGV, HIV and LGV-HIV co-infection

The force of infection is the rate at which the susceptible become infected and it is dependent on the prevalence of infection in the population. It operates through formation and dissolution of partnerships (adjusted contact rate is defined by c^*_{ijpq}) which are formed based on mixing matrix ρ_{ijpq} which assigns mixing based on perceived HIV-status *p* and activity class *i* forming partnerships with individuals of perceived HIV-status *q* and activity class *j*. β_{lgv} is the LGV transmission probability per partnership, and α is the HIV-cofactor for increased transmission of LGV from people who are HIV-infected (which is not used in this study and is set to 1).

The force of infection for LGV is calculated as:

(9)

$$\lambda_{il} = \alpha\beta_{lgv} \sum_{jq} \left(c^{*}_{ijpq} \rho_{ijpq} \frac{I_{jm}}{\sum_{q} N_{jq}} \right)$$

When *l*=2…4 (LGV susceptible is HIV-infected) and for HIV uninfected (*l*=1):

(10)

$$\lambda_{il} = \sum_{j} \left(c^{*}_{ijpq} \rho_{ijpq} \beta_{lgv} \frac{I_{j1}}{\sum N_{jq} = 1} + \sum_{m=2}^{m=4} c^{*}_{ijpq} \rho_{ijpq} \alpha\beta_{lgv} \left(1 - \beta_{hiv_n}\right) \frac{I_{jm}}{\sum N_{jq}} \right)$$

This is includes the force of infection from HIV uninfected LGV-infected and HIV and LGV-infected with the latter part using the probability that one acquires LGV but not HIV: β_{lgv} (1-$\beta_{hiv(n)}$). The force of infection for HIV operates on the same principles (defined in equations (**11)**-(**13)**). $\beta_{hiv(n)}$ is the HIV transmission probability per partnership depending on the HIV stage (unaware, aware, on ART; *l*=2…4). φ is the STI cofactor effect when an HIV-infected individual also has LGV (which is not used in this study and it is set to 1). For LGV susceptible (X_i1) the force of infection for HIV is:

(11)

$$\kappa_{i(X)} = \sum_{j} \sum_{m=2}^{m=4} \left(c^{*}_{ijpq} \rho_{ijpq} \beta_{hiv_n} \frac{X_{jm} + \varphi(1 - \beta_{lgv}) I_{jm}}{\sum_{q} N_{jq}} \right)$$

The force of infection for LGV infected (I_i1) HIV-negative is:

(12)

$$\kappa_{i(I)} = \sum_{j} \sum_{m=2}^{m=4} \left(c^{*}_{ijpq} \rho_{ijpq} \beta_{hiv_n} \frac{X_{jm} + \varphi I_{jm}}{\sum_{q} N_{jq}} \right)$$

The force of infection for HIV-LGV co-transmission is:

(13)

$$\gamma_i = \sum_{m=2}^{m=4} \left(c^{*}_{ijpq} \rho_{ijpq} \beta_{lgv} \beta_{hiv_n} \alpha\varphi \frac{I_{jm}}{\sum_{q} N_{jq}} \right)$$

The population in group *il* is defined as:

(14)

$$N_{il} = S_{il} + I_{il}$$

Where S_{il} is X_{il} or Z_{il}.. The population which perceived themselves as HIV-negative is defined as:

(15)

$$N_{i,p=0} = X_{i1} + I_{i1} + Z_{i2} + I_{i2}$$

And similarly the population who are diagnosed with HIV are defined as:

(16)

$$N_{i,p=1} = Z_{i3} + I_{i3} + Z_{i4} + I_{i4}$$

5.4.6 Mixing matrix

The mixing matrix describes how the proportion of partnerships is to be shared between the groups in the model. The mixing in the model is composed of mixing by perceived HIV status and by sexual activity, with the latter nested in the former. Similar approach was used in a paper by (White, Ward, Cassell, *et al.*, 2005). The mixing matrix is calculated as:

(17)

$$\rho_{ijpq} = \left[\varepsilon_1 \frac{\sum_j c_{jq} N_{jq}}{\sum_j \sum_q c_{jq} N_{jq}} + (1-\varepsilon_1)\delta_{pq} \right] * \left[\varepsilon_2 \frac{\sum_q c_{jq} N_{jq}}{\sum_j c_{jq} N_{jq}} + (1-\varepsilon_2)\delta_{ij} \right]$$

In the mixing matrix ε_1 is the parameter defining mixing for apparent HIV status and ε_2 for activity class. δ_{pq} is the Kronecker delta for the apparent serostatus (δ_{pq}=1 if p=q and δ_{pq}=0 if $p \neq q$) and δ_{ij} for the activity class (δ_{ij}=1 if i=j; δ_{ij}=0 if $i \neq j$). When either mixing parameter is 0 the mixing is fully assortative for that category and when parameter it is 1 the mixing is proportionate (based on available partnerships). This results in mixing matrix structure where l=1,2 and l=3,4 have an identical mixing matrix.

The identity matrix for HIV status is:

	Negative / unaware	Diagnosed / on ART
Negative / unaware	1	0
Diagnosed / on ART	0	1

And for activity class:

	High-activity	Medium-activity	Low-activity
High-activity	1	0	0
Medium-activity	0	1	0
Low-activity	0	0	1

The mixing matrix follows the principle of

(18)

$$\sum_{jq} \rho_{ijpq} = 1$$

5.4.7 Partner change rate

The sexual partnerships between groups have to balance in a closed population (Garnett and Anderson, 1994). Here partnerships are counted by activity class *i* and perceived HIV status *p* as defined in the mixing matrix:

(19)

$$N_{ip}\ c^{*}_{ijpq}\ \rho_{ijpq} = N_{jq}\ c^{*}_{jiqp}\ \rho_{jiqp}$$

To conform to the constraints of equation (19) the partnerships were balanced according to formulae described by Garnett and Anderson (1994, p.167-168), and this adjusted partner change rate is used in the force of infection:

(20)

$$c^{*}_{ijpq} = c_i \left(\frac{c_j\ N_{jq}\rho_{jiqp}}{c_i\ N_{ip}\rho_{ijpq}} \right)^{\theta}$$

And for the opposite contact rate:

(21)

$$c^{*}_{jiqp} = c_j \left(\frac{c_j\ N_{jq}\rho_{jiqp}}{c_i\ N_{ip}\rho_{ijpq}} \right)^{-(1-\theta)}$$

Above, θ is a parameter controlling for the level of compromise between the two groups across which a partnership is formed. θ is set to 0.5 so that the compromise is equal.

However in this MSM population the partnerships are automatically balanced (in other words $c^{*}_{ijpq}=c_i$). If I were to introduce strategic positioning in the model later on (with insertive and receptive roles being taken in partnerships) this would require the balancing equations.

5.4.8 Initiation of ART

On average 71% (range 69-77%) of MSM in the UK were receiving ART between 2000-2009 (Public Health England, 2010c). In a previous study, with data from 1997-2001 (Murphy, Charlett, Jordan, *et al.*, 2004) they

presented a bar chart of proportion of MSM receiving ART which was somewhere between 60-70% based on their figure.

So it would seem the proportion of HIV-diagnosed MSM on ART is increasing at a slow rate. To follow this relatively stable pattern, I am constraining the parameter of initiation of ART (ω_2) so that proportion of HIV-diagnosed (*I*=3,4) on ART is 70%. This is done in the following manner:

(22)

$$\frac{N_4}{N_3+N_4}=0.7$$

Looking at the movement between compartments, the number of individuals in each compartment, N_3 and N_4, can be calculated at each moment:

(23)

$$N_3 = N_2\omega_1 + I_2\sigma' - N_3\mu - N_3\omega_2$$

(24)

$$N_4 = N_3\omega_2 - N_4\mu$$

Therefore equation (22) can be written as:

(25)

$$\frac{N_3\omega_2 - N_4\mu}{N_2\omega_1 + I_2\sigma' - N_3\mu - N_3\omega_{2+} N_3\omega_2 - N_4\mu}=0.7$$

Rearranging this we get ω_2:

(26)

$$\omega_2 = \frac{0.7(N_2\omega_1 + I_2\sigma' - N_3\mu - N_4\mu) + N_4\mu}{N_3}$$

To avoid N_3=0 occurring in the simulations, I added additional constraints: If any N_{i3}<1, ω_2=0 so that if any N_3 activity class (i) state becomes depleted of people, ART initiation rate stops (ω_2=0). For simplicity men on ART who are lost to follow-up is not considered. Also, non-adherence to treatment regimen is not considered.

5.4.9 Latin hypercube sampling

Latin hypercube sampling (LHS) is a manner of efficiently sampling combinations of parameters from multi-dimensional parameter space. In this study LHS is used for exploring the uncertainty in the parameters. As described in Blower and Dowlatabadi (1994) LHS uses probability distributions that are assigned *a priori* to each parameter sampled. Uniform distributions were used in this study to cover the whole parameter range assigned (this is the least informative way of sampling from a parameter space). The probability distributions are then stratified into as many equal-sized (equi-probable) samples as there are simulations in the LHS. These are then randomised and paired with other parameters that are treated the same way. Thus LHS simulations explore the whole (pre-assigned) parameter space using each parameter estimate once and randomly allocating the parameter combinations.

The Matlab code for LHS that was adapted for this study was obtained from Dr Juan Vesga with the original script written by Dr Tim Hallett.

5.4.10 HIV equilibrium prevalence

MSM communities have changed in respect to sexual behaviour and demographic structure over the duration of the HIV epidemic with changes in partner number and types of sex, as well as changes in where sex partners are met (Hart & Elford, 2010; Fenton & Imrie, 2005; Davies, Reid & Weatherburn, 2002)[32]. Reproducing this would be challenging. Based on available evidence HIV incidence in MSM in the UK has not shown significant changes despite changes in diagnoses from 1995 to 2001 by Murphy, Charlett, Jordan, *et al.*, (2004) and from 2001 to 2010 by Birrell, Gill, Delpech, *et al.*, (2013). Both studies, using back-calculation approach, conclude that there is no evidence of HIV incidence decline in MSM, and neither found a significant increase in HIV incidence. However, in a recently published modelling paper HIV incidence in MSM in the United Kingdom is estimated to be on the increase (Phillips, Cambiano, Nakagawa, *et al.*, 2013).

[32] There is theoretical and empirical evidence for increased risk behaviour in the HAART era; theoretically this has been explained by reduced death rate in the high-activity individuals causing a replenished availability of sex partners in one's sexual environment (reviewed by Davies, Reid & Weatherburn, 2002; changing network structures suggested by Boily, Godin, Hogben, et al., 2005).

Therefore I decided to introduce LGV when HIV is in equilibrium; an assumption of the model, which offers advantages: as the simulations are in equilibrium, the disease dynamics are dependent on the force of infection and it is independent of initial conditions (size of the MSM population, time when HIV was introduced into the population, behaviour change over the course of the epidemic, introduction of ART and so on). This assumption therefore has the advantage of not having to make as many assumptions of the course of the HIV epidemic in MSM.

It is also worth noting that I am interested in the conditions that gave rise to LGV. The HIV epidemic is one of the primary features of this re-emergence, but it is not the primary focus of the model. Furthermore, as HIV has been circulating in the population longer than LGV, we might be more confident in behavioural parameters that produce the desired HIV prevalence range than in behavioural parameters creating LGV epidemic. Therefore I considered it easier for model interpretation to first have a desired HIV-epidemic occurring in the model – as we might expect in real life where HIV has been established in MSM for longer than LGV – and then explore how these conditions are affecting the LGV epidemic. The next section explains the technical aspects of estimating when HIV has reached equilibrium prevalence.

In order to determine when HIV equilibrium is reached (for the overall population prevalence), a criterion was needed, as numerical solution of ODEs mean that equilibrium is approached asymptotically but never exactly reached. I created equilibrium tolerance criterion T defined in equation (27) which compares HIV prevalence at yearly basis ($P_{hiv,Y}$) to the prevalence two years previously ($P_{hiv,Y-2}$). Both ratio and difference between the measurements were taken into account and they are combined using parameter X, which was set to 0.5 so that half of equilibrium tolerance criteria were derived from relative difference and the other half from absolute difference between any two years. Equilibrium tolerance criteria was set to 0.0001 in order for it to be specific enough for very slowly evolving epidemics, and for the same reason a 7-year “burn-in” period was assigned before the equation started measuring prevalence. Seven years was selected based on prior sensitivity analysis where different timeframes were tested to see which was able to capture the simulations where HIV epidemic progressed very slowly.

(27)

$$T = \chi \left| \frac{P_{hiv,Y} - P_{hiv,Y-2}}{P_{hiv,Y-2}} \right| + (1-\chi) \left| P_{hiv,Y} - P_{hiv,Y-2} \right|$$

For analysis purposes the HIV prevalence at equilibrium was included if it was between 8-30% and these were categorised into four categories: 8≤x≤15%, 15<x≤20%, 20<x≤25% and 25x≤30%. The resulting range of HIV prevalence can be thought of as separate scenarios representing different subpopulations of MSM.

5.4.11 LGV introduction to HIV equilibrium and selection of model simulations

For each of the parameter sets that were accepted on the basis of their equilibrium HIV prevalence a second round of simulations was performed where LGV was introduced in the model. Sets of parameters for LGV were also sampled by LHS.

The LGV recovery rates (σ, σ', σ'') have the same range and distribution in the Latin Hypercube Sampling. However, the data indicates that HIV-negative men experience longer duration of symptoms before seeking care for LGV (so that σ>σ', based on data-analysis in chapter 2), and we know it is likely that someone with an undiagnosed HIV infection who is treated for LGV will be tested for HIV rather than recover from LGV without HIV diagnosis (so I assume that σ'>σ''). Therefore a posterior control was placed to parameter estimates sampled by LHS and a subsample of the LHS parameter tables was selected following the rule: σ>σ'>σ''. From this LHS subsample 30 LGV parameter combinations are drawn and these are applied to each of the accepted HIV runs.

In a deterministic SIS (susceptible-infected-susceptible) system, provided that infection is able to invade the model population, persistence of the infection is guaranteed (Keeling & Rohani, 2008). In the model analysis, an LGV epidemic was considered to have occurred if the number of prevalent infections at the end of the simulation was greater than the number of infections introduced, (>10 infected); i.e. reproductive number was over 1.

For comparability of the runs the endpoint prevalence of LGV was taken as the main output if LGV had reached equilibrium by the end of the simulation. This was done using the same criteria presented for HIV prevalence and formulated in equation (27) comparing the end point prevalence to prevalence two years previously (with vast majority of runs fulfilling this criteria). This is not representative of the observed trend where LGV is diagnosed in low fluctuating levels, but given LGV has been present in MSM before the surveillance began (this was demonstrated in the Netherlands (van de Laar, Koedijk, Gotz, *et al.*, 2006), the United States (Schachter & Moncada, 2005) and Switzerland (Gebhardt & Goldenberger, 2005)) it is difficult to estimate when LGV was introduced in the population. Furthermore, given a deterministic model cannot reproduce a stochastic trend, I decided that an endemic LGV prevalence is a simplification that allows a better comparison of the different scenarios to each other across the difference HIV prevalence settings. I categorised the LGV simulations, in equilibrium prevalence, into simulations where LGV prevalence in the population was low (<1%, which is a rough approximate of the GUM clinic LGV positivity in MSM where the pooled estimate of rectal LGV is 0.98% (95% CI 0.8-1.2%) (based on two studies published by Ward, Alexander, Carder, et al., (2009) and; Annan, Sullivan, Nori, et al., (2009)), this was compared to higher LGV prevalence of 1-5% and a "large epidemic" with >5% LGV prevalence.

As mentioned previously one of the LGV recovery parameters moves people from the HIV unaware compartment to HIV aware compartment due to HIV being diagnosed when they receive care for LGV. This coupled with serosorting changes the behaviour post-HIV diagnosis (and also the transmission probability for HIV) and can thus affect the HIV epidemic. The extent to which LGV introduction in the model can change the HIV-equilibrium prevalence depends on the magnitude of LGV epidemic, the proportion of HIV unaware, the rate of HIV diagnosis and how assortative mixing by perceived HIV status (this was explored by manual univariate sensitivity analysis). In reality, the small number of LGV infections relative to HIV undiagnosed infections means that the impact on HIV is small on population level. To avoid a large HIV prevalence change in the model, parameter sets were excluded if the HIV prevalence after LGV introduction changed more than 2% (arbitrary cut-off point) in comparison to HIV equilibrium prevalence. In their analyses of LGV Enhanced Surveillance data Public Health England found that very few patients were diagnosed with HIV at LGV episode with 73/908 (8.3%) of HIV-diagnosed LGV cases in the surveillance data having an HIV diagnosis date within 4

months of their LGV episode and 12/33 of episodes with linkage to Recent Infection Testing Algorithm (RITA) were considered incident HIV (Dr Gwenda Hughes, PHE, personal communication).

The resulting runs are analysed according to their pre-assigned HIV prevalence categories and LGV prevalence categories. All model runs which had reached HIV equilibrium, were between 8-30% HIV prevalence, HIV prevalence changed <2% after LGV introduction and had an LGV epidemic which had reached equilibrium prevalence were accepted for analysis.

To assess the impact of the underlying HIV prevalence into LGV re-emergence all the results are stratified according to the underlying HIV prevalence in the population. I will look at how the prevalence of HIV and LGV is distributed between the sexual activity groups, and how HIV prevalence in LGV cases varies according to LGV epidemic size. I will then address the second aim of how serosorting affects the concentration of the two infections by describing how serosorting and activity mixing vary depending on the HIV prevalence and LGV epidemic size. I will end the results by describing sensitivity of the serosorting parameter to population heterogeneity in sexual activity and to proportion of HIV that is undiagnosed.

5.5 Results

5.5.1 Simulations selected

The primary purpose of this model was to explore the underlying processes that create the overlap between LGV and HIV epidemics. I will first describe how the model simulations were grouped:

To represent different HIV prevalence settings where MSM form partnerships I wanted to investigate a wide range of model simulations. For the analyses I selected simulations that were above 8% and below 30% equilibrium HIV prevalence. Fifteen-thousand model simulations using LHS sampling resulted in 880 runs that met this criterion. These were divided into 4 HIV prevalence categories[33]: 8-15% (number of simulations in this category: n=94), 15-20% (n=174), 20-25% (n=236) and 25-30% (n=376). The categorisation is useful as it allows

[33] Using total HIV prevalence instead of diagnosed HIV prevalence; unless separately mentioned the total HIV prevalence is used in all outputs

us to explore possible patterns of LGV re-emergence that could be due to underlying HIV prevalence in the population, and the scenarios can be broadly thought of representing different HIV prevalence senarios with 8-15% being the closest to the general MSM population prevalence, 15-20% approximating the prevalence at the gay scene (such as bars and clubs), and the two higher prevalence groups, 20-25% and 25-30%, are representing the more high-risk scene such as saunas, sex parties and other locations where HIV prevalence is likely to be much higher than in the general population (both behavioural and biological parameters were varied in the simulations so the differences between the settings are a combinations of these).

LGV was introduced into each simulation once HIV had reached equilibrium. For the four LGV parameters, 30 LHS samples were drawn and applied to each set of parameters selected on the basis of HIV equilibrium prevalence, resulting in 26,400 (30*880) model runs. Of these, 8185 (31.0%) resulted in an LGV epidemic that fulfilled the predefined conditions (LGV epidemic had occurred and it was in equilibrium by the end of the simulation, and the change in HIV equilibrium prevalence was <2%). These were distributed with 759 simulations for the 8-15% HIV-setting, 1418 for the 15-20% HIV scenario, 2318 for the 20-25% HIV scenario and 3690 for the 25-30% HIV scenario.

The simulations were further stratified by the size of the LGV epidemic with <1%, 1-5% and >5% LGV prevalence categories. There was a slight decrease in the occurrence of a low level LGV epidemic as the population HIV prevalence increased with 32.3%, 25.5%, 19.8% and 17.3% of all LGV epidemics having <1% LGV prevalence in 8-15%, 15-20%, 20-25% and 25-30% HIV scenarios, respectively. The inverse was true for the high level LGV epidemic with 6.3%, 18.9%, 14.6%, 26.8% of LGV epidemics having >5% prevalence in the different HIV prevalence scenarios (1-5% LGV epidemic size was relatively equally distributed with 55.6-65.6% of all simulations falling into this category).

5.5.2 Underlying HIV prevalence and prevalence of HIV in LGV infected

In this section I will look at the concentration of HIV in the population and among LGV infected, followed by description of HIV and LGV prevalence by sexual activity group. In Figure 2 we have a scatter plot of all the accepted model simulations, stratified by the different HIV prevalence scenarios (which are assigned different colours), plotting the equilibrium HIV prevalence against the HIV prevalence in the LGV cases. This shows a

high HIV prevalence in LGV cases across all HIV-settings. The lines initiating from the origin illustrate where the HIV prevalence of LGV cases would be expected to be if it were the same, twice as much, thrice as much and five times as much as the population HIV prevalence. We see that under no model simulation is the HIV prevalence in LGV as low as the population prevalence, and vast majority of the runs have the HIV prevalence of LGV cases at over twice that of the population prevalence.

This is due to both infections concentrating in the highest activity group as demonstrated in the following figures. Figure 3 displays box-and-whiskers plots of HIV prevalence by sexual activity group in the different settings (with black line displaying the median of all model simulations in that category, box represents the interquartile range around the median and the "whiskers" display the total simulation range with circles representing outlier simulations). In all HIV prevalence scenarios high-activity group has over 80% HIV prevalence (median of model simulations) with a stable trend over the settings whilst the biggest change with increasing underlying HIV prevalence is seen in the medium-activity group with an increasing HIV prevalence as the population prevalence increases.

Then looking at the LGV prevalence by sexual activity group, in Figure 4, the high-activity group has the highest LGV prevalence (median between 34% and 37% LGV prevalence). This is stable across the HIV scenarios as LGV and HIV do not interact biologically in the model. Furthermore when we look at HIV prevalence in LGV cases stratified by LGV epidemic size (Figure 5), there is an overall high HIV prevalence across the LGV epidemics and a trend with higher HIV prevalence in the lower the LGV epidemic size. For comparison the HIV prevalence in LGV cases in the LGV Enhanced Surveillance dataset is 81% among MSM included in the data-analysis in chapter 2.

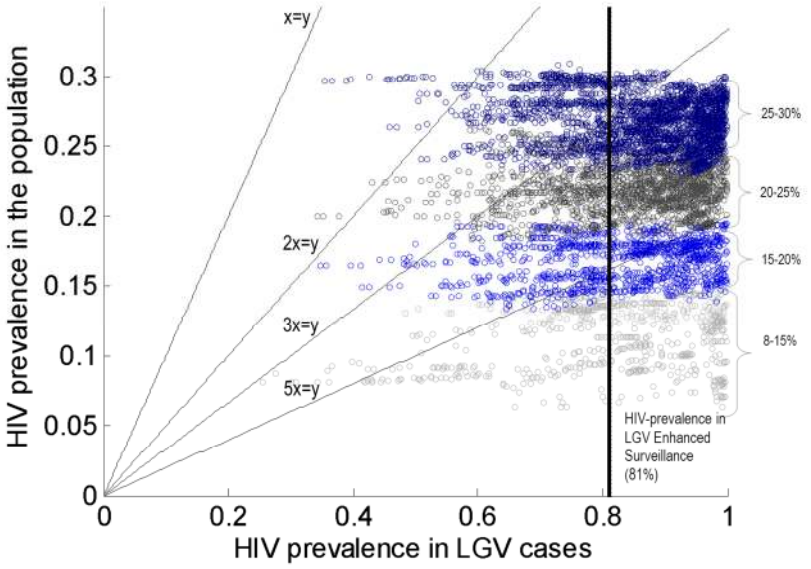


Figure 2. Scatter plot of LGV simulations contrasting HIV prevalence in the population (y-axis) to HIV prevalence in LGV cases.

Figure legend: The HIV prevalence in the population is on the y-axis and the model simulations are depicted with different colours in the scatter plot according to the HIV prevalence scenario they were grouped in. The lines compare the HIV prevalence in LGV cases to HIV prevalence in the population to illustrate where you would expect the HIV prevalence in LGV cases to be if it were the same (x=y) as the population HIV prevalence, followed by two (2x=y), three (3x=y) and five (5x=y) times the HIV prevalence in LGV cases compared to the population prevalence. The vertical line shows the overall average prevalence of HIV in MSM in LGV Enhanced Surveillance data.

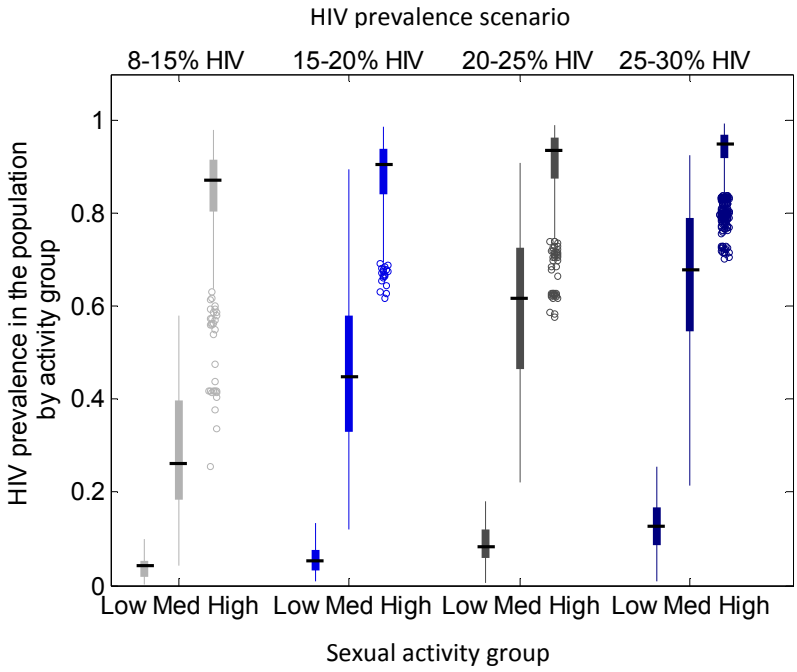


Figure 3. Box plots of HIV prevalence by activity group (low, medium and high sexual activity groups) across the different HIV prevalence settings.

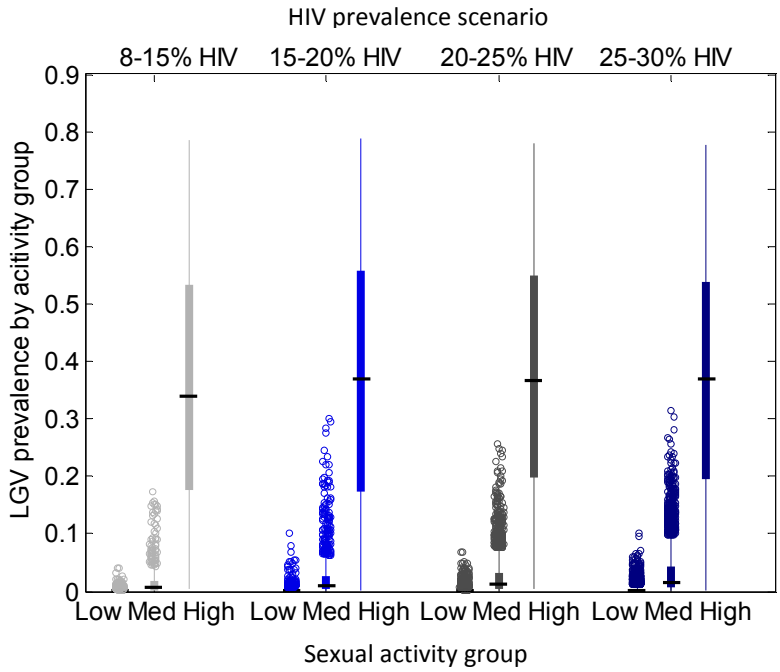


Figure 4. Box plots of LGV prevalence by activity group (low, medium and high sexual activity groups) across the different HIV prevalence settings.

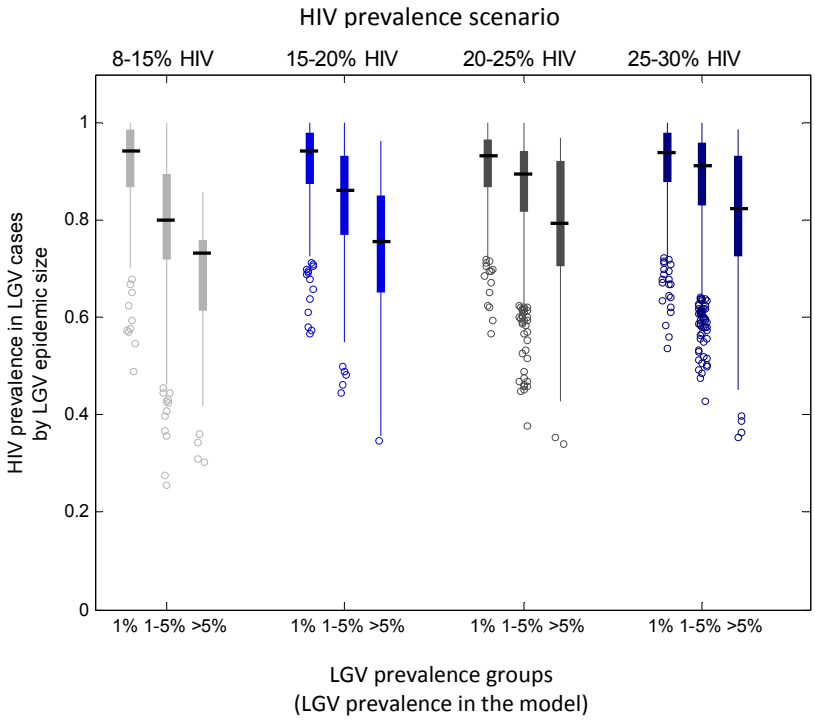


Figure 5. Box plots of HIV prevalence in LGV cases stratified by the HIV scenario and the size of the LGV epidemic (x-axis)

5.5.3 Influence of mixing patterns on HIV and LGV

The previous section examined the overlap of HIV and LGV based on sexual activity class and LGV epidemic size. This section will look at how different levels of serosorting are associated with HIV prevalence and LGV epidemic size.

Figure 6 presents how the serosorting parameter is distributed across the different HIV categories using parameter values from the accepted simulations[34]. When the parameter is 0 the population mixes purely like-with-like for perceived HIV status, and when the parameter is 1 the mixing is proportionate, that is according to available partnerships rather than preference. This is presented on y-axis showing an increasing distribution of possible mixing values as the HIV prevalence range in the population increases. A similar trend is observed

[34] For this figure this was done using parameter estimates from the HIV LHS (instead of using the second LGV LHS simulation parameters) to reduce the background noise as all the behavioural parameters were multiplied by 30 for the LGV LHS simulation. The effect of this is seen in Figure 7 and Figure 8 which use parameter values from LGV LHS.

for sexual activity mixing but the range of possible parameter are more relaxed and overall activity mixing, though assortative, would seem to have less of an influence for HIV prevalence in this model.

I also looked at mixing parameters in relation to LGV epidemic size, as presented for serosorting in Figure 7 and Figure 8 for mixing by sexual activity class (the parameter range are taken from the accepted LGV runs). Overall, more assortative mixing seems to be associated with lower LGV epidemic size, and this is most pronounced in the 8-15% HIV prevalence category. Overall this trend is not consistent with the large LGV epidemic (>5% LGV prevalence) behaving somewhat differently to the smaller LGV epidemics (where serosorting is associated with the LGV epidemic size). For activity mixing, there is more variation in mixing patterns between LGV and HIV epidemic settings.

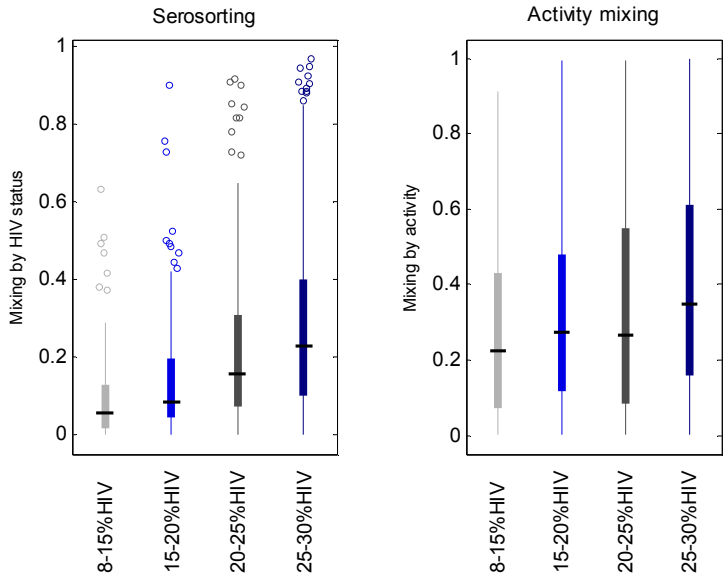


Figure 6 Results from HIV simulations stratified by HIV prevalence at equilibrium (x-axis). Parameters controlling for mixing (y-axes) by HIV status (serosorting) and mixing by activity class are presented with boxplots with 0 for fully assortative mixing and 1 for proportionate mixing.

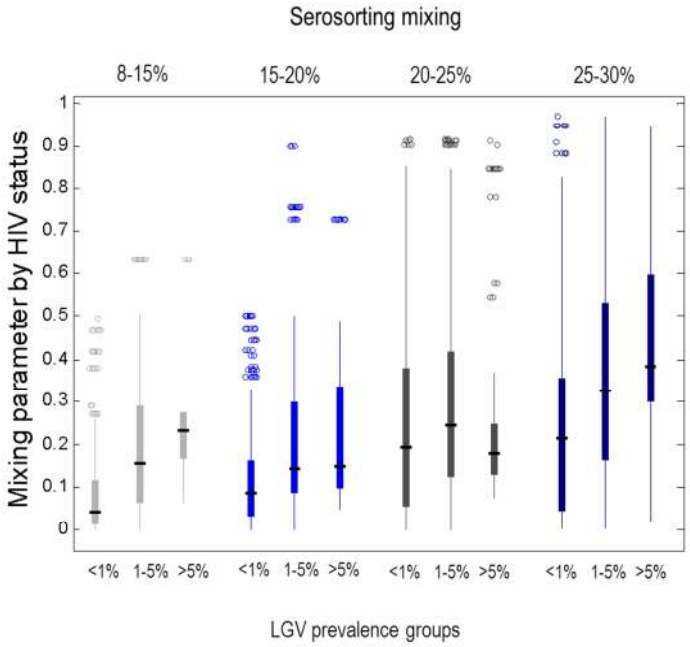


Figure 7. Results from LGV simulations stratified by HIV prevalence (top of the figure) at equilibrium and LGV prevalence in the population (x-axis). Parameter distribution for mixing by HIV status (serosorting) is presented on y-axis (with 0 for fully assortative mixing and 1 for proportionate) mixing.

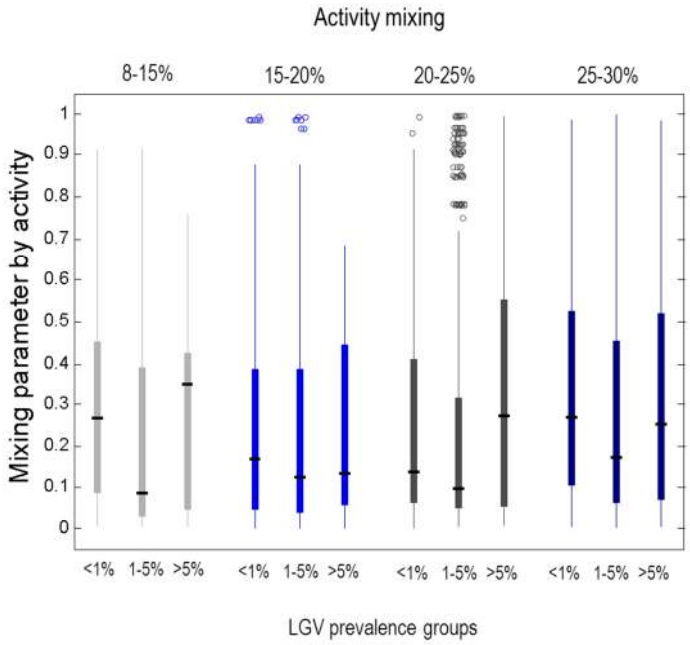


Figure 8. Results from LGV simulations stratified by HIV prevalence (top of the figure) at equilibrium and LGV prevalence in the population (x-axis). Parameter distribution controlling for mixing by sexual activity group is presented on y-axis (with 0 for fully assortative mixing and 1 for proportionate mixing).

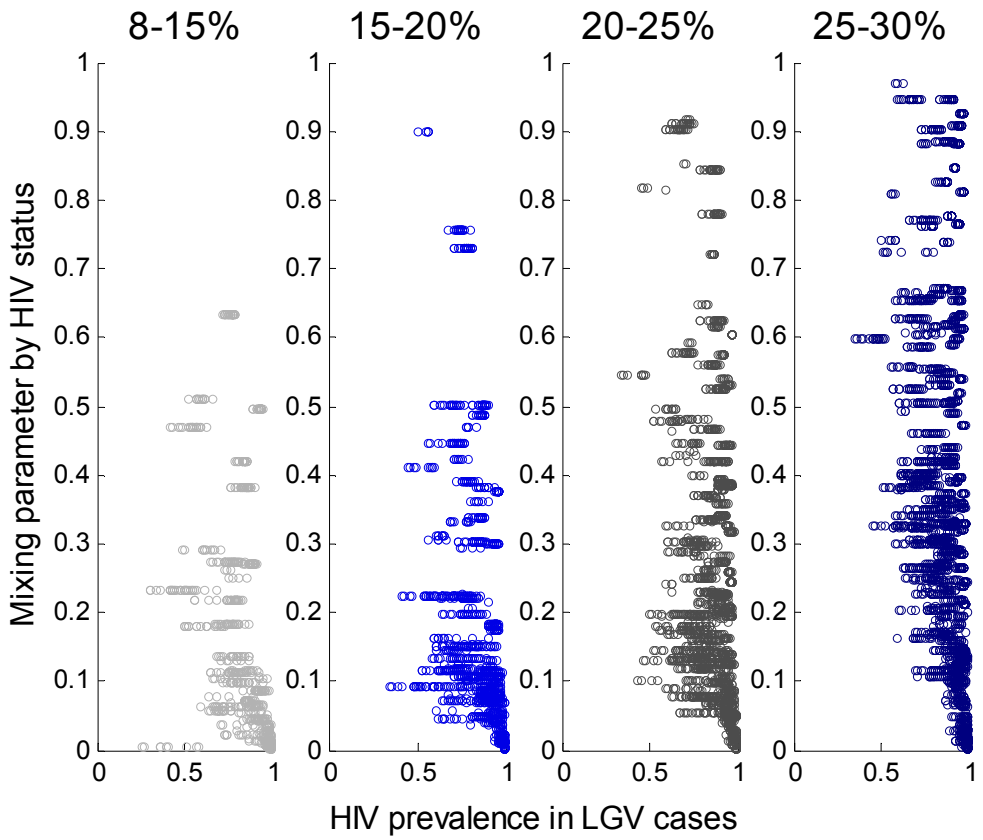


Figure 9.Scatter plots of LGV simulations comparing the HIV prevalence in LGV cases (x-axis) to mixing by HIV status (where mixing is fully assortative for serosorting when the parameter is 0 and proportionate when the parameter is 1).

Finally I looked at the impact of serosorting on HIV prevalence in LGV cases. Figure 9 shows scatter plots of the simulations with the mixing parameter for serosorting on y-axis and the HIV prevalence in LGV cases on x-axis. HIV prevalence in LGV cases is 100% when serosorting is 0 (all mix with those of the same perceived HIV status), and when serosorting parameter increases the variation in HIV prevalence in LGV cases also increases. The horizontal lines formed by the simulations result from the LGV Latin Hypercube Sampling being run 30 times on each of the accepted HIV simulations. Thus the behavioural and HIV-related parameters are fixed for these simulations and the variation in HIV prevalence in LGV cases is due to differences in LGV-related parameters. There seems to be an increasing variation in the HIV prevalence of LGV cases when serosorting becomes less assortative.

5.5.4 Sensitivity of the model to population heterogeneity and undiagnosed HIV

This section aims to address some known uncertainties in population sexual behaviour, as discussed in the methods under sexual risk behaviour considerations for the model (5.4.3), and to investigate the relationship between undiagnosed HIV and serosorting.

Given the epidemiological importance of the high-activity group to the transmission of both HIV and LGV, I wanted to explore how the highest risk group (their size and activity, both of which are varied in the Latin Hypercube Simulation) contributed to the mixing patterns. As a measure of the dispersion I compared the number of partnerships generated by the high-activity group to the mean number of partnerships generated by the model population (difference in the two calculated for each run). This was plotted against the serosorting and activity mixing parameters for the low LGV epidemic (with <1% LGV prevalence), and presented in Figure 10. For serosorting, it would seem that the more dispersion (the more the high-activity group partnerships differ from the population mean) the more assortative the mixing. For activity mixing, this is not evident. For comparison similar plot for 1-5% LGV prevalence epidemic is presented in appendix Figure S3.

The proportion of undiagnosed HIV-infected of all HIV-infected varied from under 10% to close to 30% in the model runs with little difference between the HIV prevalence categories. In appendix (Figure S1) I plotted the proportion of HIV-infected who are unaware against serosorting mixing parameter to test if they are correlated (as serosorting is based on perceived HIV-status), but there appears to be no visible relationship between the two in these simulations with HIV in equilibrium.

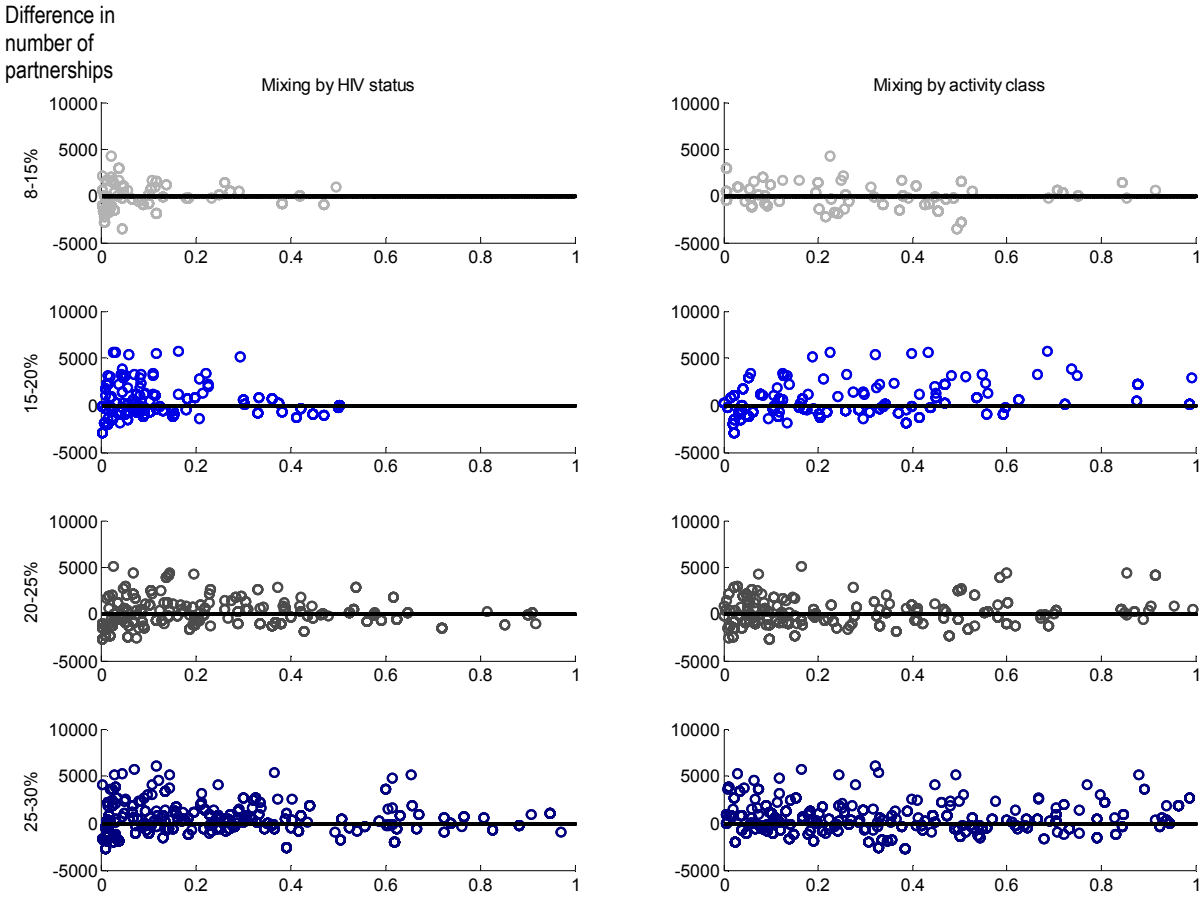


Figure 10. Scatter plots of parameter estimates for serosorting and activity mixing against the dispersion of number of high-activity partnerships compared to the population mean number of partnerships. Results are for <1% LGV prevalence epidemics.

Figure legend: On the y-axis is the difference between numbers of partnerships generated by the high-activity group compared to mean number of partnerships generated by the activity classes in the model population. On the x-axis we have the mixing parameter for serosorting (panels on the left) and for activity mixing (panels on the right). The horizontal line at zero represents a situation where high-activity group partnerships are the same as the mean number of partnerships generated by the population (population heterogeneity low).

5.6 Discussion

The primary purpose of this study was to explore the underlying behavioural processes that create the overlap between LGV and HIV epidemics. These were determined through heterogeneity in risk behaviour with three sexual activity groups, and mixing by activity group and by HIV status. All of these were varied between the runs along with the natural history parameters for LGV and HIV. A deterministic model of HIV-progression was created which was coupled with an SIS model of LGV. Parameter sets generated by Latin Hypercube Sampling resulting in different equilibrium HIV prevalence were selected for equilibrium HIV prevalence between 8-30%. LGV was then introduced through a second LHS for LGV parameters and the resultant model simulations were examined when both infections were in equilibrium prevalence. Differences in HIV prevalence and LGV prevalence were compared to examine the extent of the overlap under different scenarios for LGV and HIV, and how mixing patterns for perceived HIV status and activity mixing may contribute to this.

Three primary findings emerged from the analyses of HIV-LGV mathematical model. Firstly, even with a large variation in the background parameters, LGV and HIV tend to concentrate in the same high-activity class, and this is particularly pronounced when LGV prevalence is low (<1%). This results in a high HIV prevalence in LGV cases, across the different HIV prevalence scenarios. Secondly, there is an association between more assortative serosorting and lower HIV prevalence on population level as well as assortative serosorting and low LGV prevalence. The distribution of possible serosorting parameter values increases as the underlying HIV prevalence in the population becomes larger. Lastly, in the sensitivity analysis I found that serosorting was also related to the population heterogeneity in sexual activity with more assortative serosorting associated with greater dispersion in sexual activity between the activity groups.

This chapter demonstrates that heterogeneity in sexual risk behaviour coupled with a low-level LGV epidemic and serosorting can explain the observed phenomenon of LGV being acquired and transmitted mainly by HIV-positive individuals. This supports the earlier work, and the idea that high-risk HIV-positive MSM drive the LGV epidemic. In the light of this work, serosorting can have an impact on STI trends, and interventions should

target appropriate behaviours to mitigate this, such as more frequent/feasible STI testing to shorten the duration of infection.

5.6.1 Strengths and limitations

The natural histories of HIV and LGV are represented through as simple structure as was feasible. This helped me to understand the model behaviour and underlying processes through a wide range of situations, but makes the results theoretical in nature. Clear trends were apparent when other behavioural and biological parameters were varied simultaneously, which strengthens the reliability of the findings.

Given the simplicity of the model, the importance of serosorting might be an overestimate and other factors – changes in HIV-infectivity and general changes in risk behaviour over the duration of HIV infection – may have a bigger role in determining the HIV prevalence in a population than serosorting alone.

A study that looked at the impact of undiagnosed HIV on the preventive value of serosorting, concluded that when the proportion of undiagnosed HIV exceeds 20%, serosorting increases HIV transmission (Wilson, Regan, Heymer, *et al.*, 2010). In this study there was no direct association between level of serosorting and proportion of undiagnosed HIV. In Wilson's study the estimates were point-estimates and epidemic phase of HIV was not considered. In this study HIV was in equilibrium in all the model runs, and the effect between serosorting and proportion undiagnosed might occur earlier in the growth phase. Previous empirical studies have estimated the proportion of undiagnosed HIV to be as high as 40% in venue sample in London (Dodds, Johnson, Parry, *et al.*, 2007), whilst PHE estimates the proportion undiagnosed to be 26% (Delpech, 2012).

As was noted compartmental deterministic models of STI transmission are affected by those in the model with the highest sexual activity, the group's size and their partner change rate. The effect of sexual heterogeneity was explored in the sensitivity analysis of the model. The equilibrium state has many benefits for this analysis, but given the potentially increasing incidence of HIV in MSM as indicated by a recent mathematical model (Phillips, Cambiano, Nakagawa, *et al.*, 2013), there may be features in the increasing incidence that is also of interest for re-emerging STIs; such as increase in the size of the population infected with HIV and potential changes in sexual behaviour. Another feature of the model is that loss to follow-up on those on ART was not

considered. Also, the emphasis of this chapter was on serosorting, which has been modelled as assortative mixing based on your own and your partner's perceived HIV status. As the previous chapter demonstrated seroadaptive behaviours are more complicated than this and also situation and partnerships specific. This is an important limitation to keep in mind and given the interest and data-generation around seroadaptive behaviours it is also a possible avenue for further mathematical modelling.

5.6.2 Future research

The intersecting epidemics of LGV and HIV was modelled purely from a behavioural point of view, and this work suggests that behaviour can be a major explanatory factor for the clustering of LGV and HIV in the same sub-populations. As an ulcerative STI, LGV is likely to increase the transmission of HIV and it is an important consideration which could be explored further.

The aim of the model was not to reproduce precisely the observed LGV epidemic, and there are important features of LGV re-emergence that remain unanswered and may have a large contribution to its transmission: potential asymptomatic/pre-symptomatic reservoir of LGV, potential differences in transmission depending on sexual act and how the natural history of LGV might differ between individuals (based on infection site, HIV-positivity or some other factor). However, given the scarcity of data around these topics, exploring this through a mathematical modelling would create a large range of scenarios from which it would be hard to select the ones that are relevant. If we wanted to make LGV epidemic more realistic the first step would be to increase complexity in the SIS model (for example by adding asymptomatic and symptomatic states into LGV natural history) and make it stochastic. If there were data available, an individual-based model would be the best choice to capture the dense sexual networks where LGV is thought to circulate. This approach would also measure the sexual behaviour with greater precision (Cassels & Goodreau, 2011; Morris, 2010). Collecting these data would require extensive contact-tracing.

5.7 Conclusions

This chapter demonstrated that LGV and HIV tend to concentrate to the same high-activity individuals. High sexual activity coupled with serosorting and low-level LGV epidemic can further explain the high level of HIV prevalence in LGV cases as demonstrated. Many of the concepts here, regarding heterogeneity and assortative

mixing as important features of STI persistence, have been demonstrated in seminal works in STI epidemiology (most notably by Anderson & May, 1991). However, serosorting is a more recent phenomenon and this study adds to the understanding of how serosorting and high-risk behaviour can explain the re-emergence of STIs in HIV-positive MSM.

Chapter 6

Conclusions

6.1 Summary

In this chapter I will summarise the findings of this study and discuss them in relation to the strengths and weaknesses of the analyses, the context of the research and future research avenues. The chapter will conclude with a discussion of implications of the work.

The work began from the observation of a strong association between LGV and HIV. The overall aim of the thesis was to investigate reasons for the association between LGV and HIV from epidemiological and public health standpoints. I have used statistical analyses based on LGV surveillance data, literature review coupled with social epidemiological perspective and mathematical modelling as the primary tools to explore the research questions.

This thesis found that HIV-positive LGV patients report more sexual risk behaviour than HIV-negative and unknown LGV patients. They also reported a shorter duration of symptoms indicating diagnostic bias in favour of HIV-positive MSM who have increased contact with the health care system. I moved on from the individual-level risk behaviour to the social context of LGV re-emergence. Through a literature review I proposed a conceptual framework of seroadaptive behaviours in HIV-positive MSM. The role of sexual activity level and serosorting were further explored in a mathematical model which gives further support to the hypothesis that serosorting can explain the particularly high HIV prevalence in those who acquire LGV.

6.2 Introduction

LGV represents a new public health problem, and is an added burden in the control of STIs. LGV re-emergence has been characterised by atypical clinical manifestation, the need for specialised diagnostics and new surveillance systems to monitor the occurrence of the disease. Patients have been predominantly high-risk, HIV-positive MSM and LGV re-emergence has occurred along with a parallel increase in other STIs in this population.

In the next section I go over the main findings regarding the utility of LGV surveillance and risk-behaviour of LGV patients as observed in the LGV Enhanced Surveillance data. I then discuss the framework of seroadaptive behaviours in HIV-positive MSM and further explore the role of serosorting in HIV-LGV mathematical model.

6.3 LGV surveillance

LGV re-emergence has been treated as an outbreak, and LGV Enhanced Surveillance was established to monitor this. The laboratory surveillance is based on testing patients with LGV-type symptoms and their partners, and the Enhanced Surveillance was based on voluntary reporting by clinicians of confirmed LGV cases. The UK was able to establish appropriate surveillance systems soon after LGV was acknowledged, and it has described the largest outbreak of LGV; this may reflect the true outbreak size but also reflects the comprehensive surveillance. The LGV Enhanced Surveillance ended at the end of 2010, and surveillance relies now on laboratory diagnoses of LGV. The new clinical surveillance system, GUMCAD, has included LGV, but due to the time delay before a patient is confirmed to have LGV, GUMCAD-based LGV reporting is not functioning to full capacity, and cannot yet be used for surveillance purposes (Dr Gwenda Hughes, personal communication).

Most research done on LGV re-emergence has been based on surveillance data generated by the affected countries. The LGV Enhanced Surveillance dataset in the UK has been used to describe the profile of LGV cases, and has documented persistently high levels of co-infection with HIV, an overlap of LGV with hepatitis C, and reported that the profile of cases has changed little since the beginning of the surveillance (von Holsterin, Fenton & Ison, 2004; Simms, Macdonald, Ison, *et al.*, 2004; French, Ison & Macdonald, 2005; Macdonald, Ison,

Martin, *et al.*, 2005; Ward, Martin, Macdonald, *et al.*, 2007; Jebbari, Alexander, Ward, *et al.*, 2007; Savage, van de Laar, Gallay, *et al.*, 2009).

Had there been a radical change in patient profile (such as cases detected in heterosexuals or large numbers of LGV cases seen outside London), the laboratory surveillance would have detected it promptly, and the LGV Enhanced Surveillance could have given more information on the reasons. However a more thorough analysis of the data, such as one done in this study is necessarily retrospective in nature, and given the delay in Enhanced Surveillance reporting[35], real-time analysis is not feasible. From the research point of view, given that surveillance was not collected for analytic purposes, the data analyst is at the mercy of the data in terms of what analyses are appropriate – such as the problem described with repeat infections in chapter 3 – which can limit the usability and inferences drawn from the data. In this study the data were used to explore LGV-HIV association and the manner of data collection was included in the interpretation of results.

The statistical analysis done in chapter two demonstrated a longer duration of symptoms for HIV-negative/unknown, which would indicate diagnostic bias in favour of HIV-positive men. Generalizability of the results in the presence of diagnostic bias is questionable. When testing is based on clinical suspicion it depends on the positive predictive value of the clinical symptoms of the disease which determines how representative the sample is of the total population with LGV. The last case-finding exercise performed (Saxon, Hughes, Ison, *et al.*, 2013) found more asymptomatic LGV cases than seen previously (16%, or 8/49 LGV patients with clinical information remained asymptomatic until treatment), and changes in testing might be advisable in high prevalence settings (in first instance, the clinics which diagnose the largest number of LGV cases)[36]. This would cause additional costs to the clinic and diagnostic laboratory.

National Enhanced Syphilis Surveillance (NESS), and Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP) are comparative models of STI Enhanced Surveillance performed in the UK. NESS operated based on voluntary reporting similar to LGV Enhanced Surveillance, and it suffered from poor

[35] Median of 98 days, distribution in appendix for chapter 2, Figure S2.
[36] In Amsterdam LGV testing is performed on all chlamydia positive rectal and ulcer samples from MSM (de Vrieze, van Rooijen, van der Loeff, *et al.*, 2013).

reporting coverage, and differential reporting (Menon-Johansson et al. (2009) for instance demonstrated that NESS under-reported heterosexual syphilis in both men and women in London compared to confirmed syphilis episodes), on the other hand GRASP is a long-standing surveillance system which operates using systematic data-collection principles repeated yearly[37] but this type of system would be difficult to establish for emerging infections.

The case-finding exercises for LGV have been useful in describing the positivity of the infection in GUM clinic attendees as well as the changing profile of LGV cases with an increase in asymptomatic LGV cases seen in the last case-finding exercise (Annan, Sullivan, Nori, *et al.*, 2009; Ward, Alexander, Carder, *et al.*, 2009; Saxon, Hughes, Ison, *et al.*, 2013). Given LGV has remained relatively rare, having periodic sentinel enhanced surveillance might be more useful for on-going outbreaks with more detailed information collected sporadically.

6.4 Risk behaviour in those who acquire LGV

I examined the enhanced surveillance dataset with an emphasis on the profile of LGV episodes occurring in HIV-positive individuals to identify differences between HIV-positive and HIV-negative subpopulations who acquire LGV. HIV-positive LGV patients, in general, reported higher levels of risky sex, and in particular unprotected receptive anal intercourse, which is a plausible route of LGV acquisition. Also reporting fisting had a non-significant positive association, and it has been suggested as a possible risk factor (de Vries, van der Bij, Fennema, *et al.*, 2008). Based on few people reporting fisting, and given it is likely to occur alongside other sexual practices, it is possibly a marker of risk-behaviour – and potentially a network membership – in general.

There is likely to be diagnostic bias also in identification of repeat infections, but from what can be observed in the data their clinical profile further strengthens the concept of LGV being circulated in the "core of the core" with the majority of repeat infection seen in HIV-positive men, LGV being more prevalent in London (where

[37] GRASP combines laboratory and clinical data on gonococcal isolates diagnosed at collaborating laboratories. There are 7 laboratories and 9 GUM clinics in London and 17 GUM clinics and laboratories outside London. All gonococcal samples taken in July, August and September by the sentinel laboratories take part in the enhanced surveillance. GUM clinics provide behavioural data to the surveillance and reference laboratories provide data for antimicrobial susceptibility assessment (Public Health England, 2010a).

most repeaters were also located) and the significant overlap with other STIs, most notably with gonorrhoea and hepatitis C, in LGV repeaters. This indicates that LGV repeaters have a central position in sexual networks where STIs circulate.

Despite the evidence of the network position of those who acquire LGV, no part of my study measured sexual networks specifically. Therefore one of the most fundamental ideas around LGV re-emergence – that it is sustained by densely connected high-risk population which is relatively small in size – is supported by this work, but the evidence remains indirect. It is also important to remember that the statistical analyses are based on cross-sectional data, and can measure associations between two variables, but not causation. As the analyses are based on reported behaviour to clinician, social and recall biases are likely to influence results.

Multi-drug use and sex parties have been described in London club scene along with HIV-positivity and hepatitis C infection being common (Kirby & Thornber-Dunwell, 2013). In context of a shigella outbreak drug-use was described as part of disinhibition and sexual experimentation (Gilbart, Simms, Gobin, *et al.*, 2013). LGV patients reporting frequent and diverse drug-use was also demonstrated in a case-control study of risk factors for LGV (Ward, Macdonald, Ronn, *et al.*, 2011). In a subset of these men – those who acquire LGV repeatedly and have also hepatitis C (which coupled with HIV can develop into chronic hepatitis and liver cirrhosis) – the behaviour is no longer merely “high-risk” but has become more self-destructive.

Given the number of sex partners and levels of unprotected sex LGV patients report, traditional STI control methods – such as partner tracing or endorsing condom use – are less likely to be practical. For STI prevention making testing more frequent to targeted groups (such as men with hepatitis C and LGV) might work better at reducing duration of infection, when one occurs, and home test kits and other means of making frequent testing more feasible might be a way of achieving this (such as testing in venues). In Australia, sending automated messages to remind MSM of regular STI testing increased the frequency of sexual health screens and rate of STI detection in this group (Zou, Fairley, Guy, *et al.*, 2013). A more forceful alternative would be a targeted treatment to at-risk population, such as mass treatment programme to curb syphilis outbreak done in Vancouver, Canada (Rekart, Patrick, Chakraborty, *et al.*, 2003).

From the point of view of HIV-prevention LGV patients who are still HIV-negative, should be offered counselling and support as they are at particularly high risk of HIV acquisition. Pre-Exposure Prophylaxis (PrEP) might be a biomedical option as PrEP can be cost-effective if targeted to those at greatest risk of HIV-acquisition (Gomez, Borquez, Caceres, *et al.*, 2012). Half of HIV-negative men in London – who responded to Gay Men's Sexual Health Survey in 2011 – reported they would likely take daily pill of PrEP if it came to market (Aghaizu, Mercey, Copas, *et al.*, 2013)[38].

6.5 Seroadaptive behaviours in HIV-positive MSM

In order to understand the social context in which LGV has re-emerged I performed a literature review to describe seroadaptive behaviours in HIV-positive MSM from a social epidemiological perspective, and used this to develop a conceptual framework.

Structural and community factors facilitate seroadaptive behaviours in HIV-positive MSM, for example the institutionalisation of HIV criminal prosecution for sexual HIV transmission. The potential public health benefit or harm this may result in is not clear. The process and barriers for disclosure may have facilitated the formation of "micro-cultures" (as described in Adam et al. (2008) who described the social scene of sex parties) whereby separate social rules of HIV disclosure and perception of risk depend on the setting where sex partners are encountered. It was interesting to see that responsibility over a partner's health seemed to be in direct relation to the type of partnership with considerable care being described over the health of the main partner and reducing sense of responsibility when the number and anonymity of sexual encounters increases.

Seroadaptive behaviour remains a concept which encompasses a diversity of practices, and the lack of standardised definitions in the literature added to this heterogeneity. An extensive systematic review on different aspects of seroadaptive behaviours was impractical given the time limitation but the somewhat more limited literature review part of this study was useful in understanding the multitude of factors surrounding

[38] Reporting likely use of PrEP (given its availability) was associated with being under 35 years, reporting UAI with a casual partner in the last year and previous Post-Exposure Prophylaxis (PEP) use. However, the authors also noted, there is no indication how much the men knew about PrEP, its efficacy and potential side-effects, and PrEP might have also been confused with PEP by some.

the topic. The conceptual framework of seroadaptive behaviours presented can be utilised in hypothesis generation, and in design and interpretation of studies. The framework is based on literature, but before it can be considered as a theoretical model it should be validated with further data-analysis, as was done with proximate determinants framework for HIV (Lewis, Donnelly, Mare, *et al.*, 2007).

The men who are most at risk of LGV acquisition represent a minority of HIV-positive MSM. Conversely there is a much greater proportion of men who practice seroadaptive behaviours, either systematically or driven by the setting. This will also mean that risk (for acquisition and transmission of HIV and/or other STIs) of seroadaptive behaviours is not uniform across individuals and locations. It would thus be premature to determine whether public health officials should promote or dissuade seroadaptive approaches in risk management.

Furthermore, even in situations where risk of STI transmission exists between HIV-positive men who elect to have unprotected sex, this risk may be justifiable for the individuals; this creates a conflict of interest between public health and personal interests. The literature also highlighted a need for more holistic healthcare for HIV-positive MSM, which would incorporate different areas of one's life – such as mental and social wellbeing, drug and alcohol use, sexuality and stigma – rather than viewing HIV-positivity merely as management of an infection. The framework highlighted the value of the community (social norms, social and sexual networks and venues), and incorporation of this to health programmes should be considered.

6.6 Conceptual modelling of HIV and LGV

A population-level approach is adopted in the fifth chapter in which I investigated the transmission dynamics of HIV and LGV through a mathematical model. This allows a theoretical exploration of the concepts presented in the previous chapters such as the influence of underlying HIV prevalence and the effect of serosorting on the association between LGV and HIV.

When HIV has been driven to equilibrium the infection is saturated in those with most acquisition risk; then in a low-level LGV-epidemic the HIV prevalence in LGV cases tends to be high, regardless of the background HIV prevalence due to both infections being concentrated in the highest activity group. Trends were apparent when different HIV and LGV scenarios were observed: serosorting was associated with lower HIV-prevalence, low-level LGV epidemic (<1% LGV prevalence) and it was also associated with increasing heterogeneity in the population sexual activity.

To my knowledge this is the first mathematical model of LGV and as the model also incorporates HIV transmission it allows investigation of interactions between the two infections. Instead of focusing on how serosorting affects the risk of HIV infection, this work looked at introducing LGV infection into a population of MSM adding to the understanding of how serosorting is related to the re-emerging STIs.

Model structure and parameter estimates have a great influence on the model results. While adding structural complexity to models makes them more realistic, it also increases the number of parameters and model uncertainty. It is important to note that the diversity of seroadaptive behaviours in the model is reduced to UAI practiced with partners of the same perceived HIV status. Strategic positioning, HIV viral load and sexual practices other than UAI are ignored. Depending on how much these contribute to the transmission of HIV and LGV, we may be missing out on important behavioural determinants. The potential of oral sex to contribute to syphilis transmission but less so to HIV transmission is an example of a situation where behaviour favours transmission of one but not the other infection (Fenton, 2004). Mathematical modelling can offer a way of exploring the effects of different seroadaptive behaviours for infection transmission.

6.7 Work in context

To illustrate how the separate areas of research fit into transmission dynamics of LGV and HIV and wider determinants of infection a schematic illustration is presented in Figure 1. Determinants of infection are categorised into transmission dynamics, natural history of the disease, individual behaviour or the sexual network in which the behaviours are embedded, health care system managing the infection and the overall social structures which influence all of these levels.

HIV-related work is presented in blue and LGV-related work in orange with work in this study in boxes with solid lines and future work in dashed lines (future work is discussed in more detail in the next section). Chapter 4 with its conceptual framework was able to cover several levels of factors contributing to HIV transmission dynamics whilst chapter 2 and 3 focused on clinical and behavioural factors of LGV. Chapter 5 with its mathematical model explored transmission dynamics of HIV and LGV and individual sexual behaviour as contributing factors to LGV re-emergence.

The strength of the approach used in this study is that it has been able to incorporate a range of methods to evaluate determinants of LGV re-emergence. The limitation of such an approach is that it is less specific and whilst it covers many topics it does not examine any in great detail.

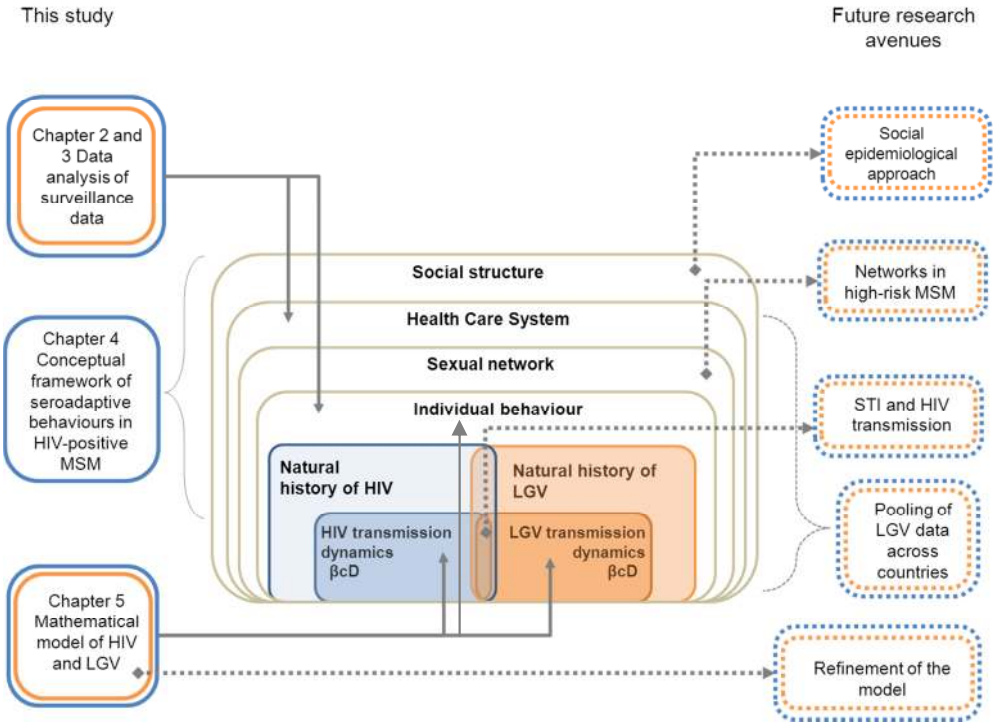


Figure 1. Schematic illustration of infectious disease framework in relation to work in this thesis and future research avenues.

6.8 Future work

LGV has established itself in the UK and Netherlands whilst many other European countries continue to report sporadic LGV cases. Given LGV networks are thought to be linked between countries – much like that described for hepatitis C (van de Laar, Pybus, Bruisten, *et al.*, 2009) – cross-country comparison would be useful, and could be done by pooling surveillance data where it has been collected. This would allow investigation of cross-country similarities in case-numbers such as the increase in case numbers in the UK in 2010 which was followed by a slight increase in cases in Barcelona in 2011 (Vargas-Leguas, Garcia de Olalla, Arando, *et al.*, 2012) as well as in the Netherlands in 2011 (de Vrieze, van Rooijen, van der Loeff, *et al.*, 2013). Pooling of the data would also increase statistical power to detect associations, and this might be a better way of addressing broad questions such as "what drives LGV epidemic".

There is an interesting parallel of findings between UK and the Netherlands regarding inguinal LGV being found more in HIV-negative men, and the few cases of inguinal LGV experienced in both countries as described in chapter 2 for the analysis of HIV-positive and –negative LGV episodes and in a paper by De Vrieze *et al.* (2013). A cross-country comparison of data could attempt to address whether this could be due to different clinical pathways between HIV-positive and –negative men with LGV (identification of LGV), indication of strategic positioning or whether the natural history of LGV could be altered by HIV status.

Incorporating the social context in which STI outbreaks occur, as done with in-depth qualitative interviews with the shigella outbreak in MSM (Gilbart, Simms, Gobin, *et al.*, 2013), can give a better understanding of the context and appropriate control measures. If we were also able to examine the sexual networks where LGV and a number of other STIs circulate, we would obtain a better understanding of the transmission dynamics. Network study would be challenging and resource-consuming in practice. Alternatively a venue-based sampling might give us insight on LGV networks similar to the study done in San Diego (Drumright & Frost, 2010).

A mathematical model has been developed as part of this thesis. Given its relatively simple structure, this model can be used to answer different mechanistic questions regarding the interplay between HIV and STIs in MSM. The potential of biological synergy between HIV and LGV can be looked at by employing the STI and HIV cofactors already incorporated in the model. LGV, which is currently nested within the HIV model as a deterministic compartmental model, could be modified into stochastic compartmental model if we wanted to better assess the stochastic nature of LGV re-emergence.

6.9 Implications

This work presents a synthesis of evidence based on incorporating individual-level statistical data-analysis, social epidemiological perspective of contextual factors and theoretical mathematical modelling. I have demonstrated that there is a strong behavioural component in explaining the association between the two infections. This cannot exclude the possibility of biological interaction but the current evidence points to behavioural, and likely network-level, differences between HIV-positive and -negative MSM to be the main driver in LGV re-emergence. This work has implications for surveillance and control of LGV, which I will summarise briefly in this section.

Currently LGV surveillance relies on the laboratory diagnoses, and tracing the epidemic further might prove difficult in the absence of LGV Enhanced Surveillance. This work has demonstrated the importance of behaviour and sexual networks for this outbreak. This would suggest that a different approach to surveillance could be used instead, for example a more detailed and qualitative investigation of cases using a sentinel model. At the moment we may be left with local level intelligence about LGV transmission in the absence of more systematic surveillance systems.

To tackle the “core within a core” where LGV and a number of other STIs circulate, a location-based intervention might be useful, as well as work with community groups and businesses to increase awareness of LGV, other re-emerging STIs and limitations of serosorting. There is a need for more holistic health care for the HIV-positive MSM and health promotion that tackles a range of issues including mental health, drug use, self-efficacy as well as disclosure and negotiation skills.

6.10 Key recommendations

The data-analysis demonstrated that HIV-negative MSM have a longer duration of symptoms before presentation to the clinic compared to HIV-diagnosed men, which would indicate that LGV testing is not as frequent in this group as in HIV-positive men. Furthermore the last LGV case-finding exercise identified a number of asymptomatic LGV cases (Saxon, Hughes, Ison, *et al.*, 2013), whilst LGV testing is currently based on clinical suspicion. We should consider modifying the testing algorithm to include men at risk of LGV even if they are asymptomatic[39]. In first instance this should include men reporting relevant high-risk behaviours in areas where LGV prevalence is elevated. Refinement of the testing algorithm can be done by analysing the case-finding data to identify factors according which to target testing.

LGV Enhanced Surveillance is no longer in operation and the surveillance currently relies on laboratory surveillance without further epidemiological knowledge of the cases. Periodic case-finding exercises can be a way of monitoring the effectiveness of current testing guidelines in detecting the cases, changes in the profile of cases and potential new behavioural risk factors which may be contributing to the transmission of LGV or other STIs (such as injecting recreational drugs which was reported during shigella outbreak in MSM (Gilbart, Simms, Gobin, *et al.*, 2013)). As HIV-negative LGV cases are at high risk of HIV acquisition, they represent a target group for enhanced HIV prevention interventions and biomedical (such as PrEP) and behavioural (such as counselling) interventions could be offered for these individuals.

In the seroadaptive behaviours chapter, we have seen that the concept covers a variety of behaviours with varying degrees of risk for infection transmission. Our understanding of what constitutes as seroadaptive behaviour among HIV-positive men, and how gay men understand and practice the behaviours is still relatively limited, and this should be explored through qualitative and quantitative studies. However, based on the literature it is clear that a subset of HIV-positive men engage in a range of risk behaviours. And we should incorporate a more holistic attitude to care of HIV-positive MSM, including acknowledging and addressing

[39] Given the low prevalence of LGV and the testing being centralised to STBRL, it is unlikely to be neither feasible nor affordable to introduce wide-spread LGV screening at this state.

issues surrounding mental health among those diagnosed with HIV. The contribution of seroadaptive behaviours to the transmission of HIV and STI can be further analysed through mathematical modelling.

6.11 Personal reflections

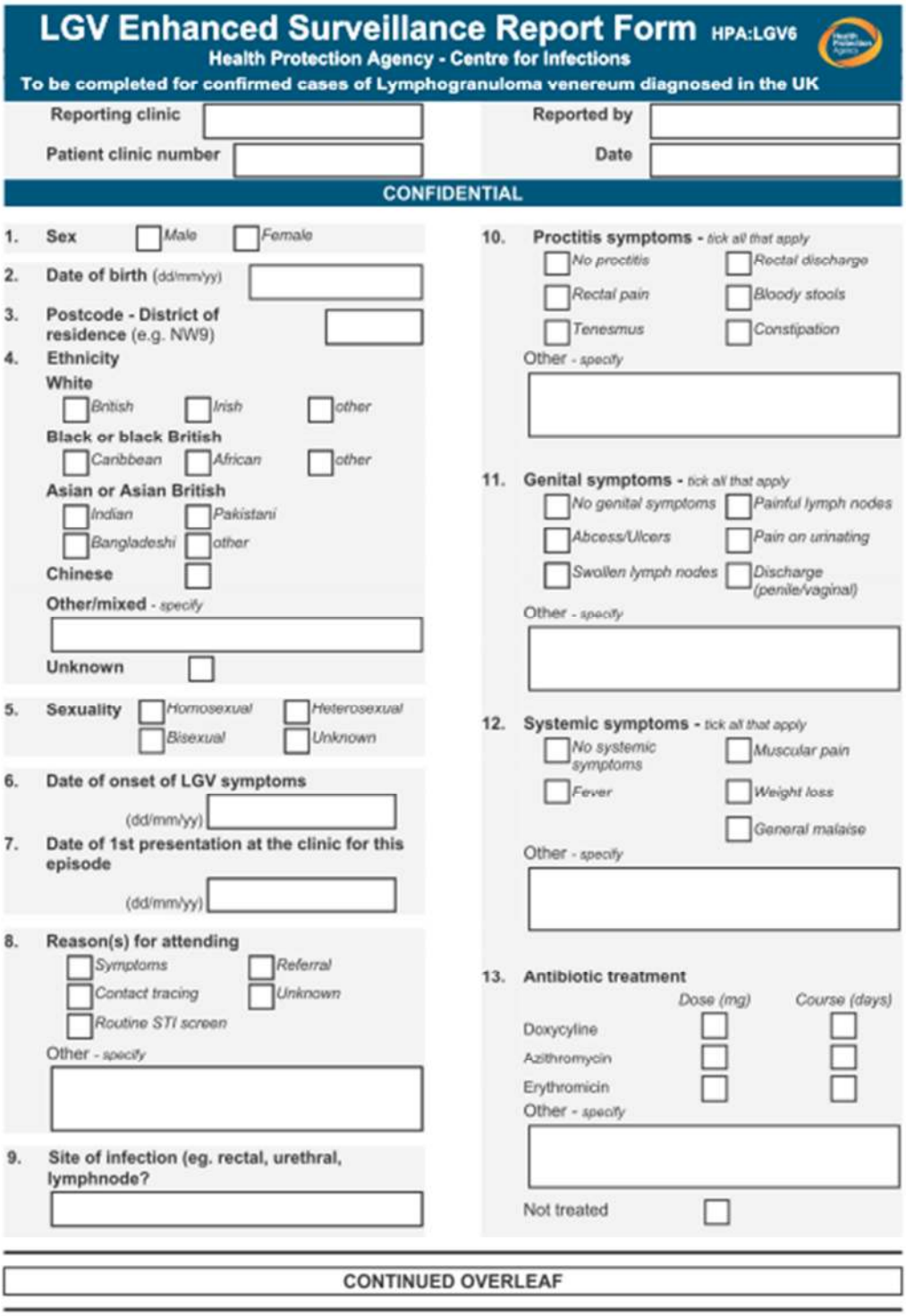
I began the PhD by focusing on LGV and trying to understand the re-emergence by looking at LGV infected in LGV Enhanced Surveillance data. Through the data-analysis and searching the literature, it became clear that the description of HIV-positive MSM with high-risk behaviour offered a rather one-dimensional explanation for the LGV re-emergence. Understanding the context of high-risk behaviours through the literature review on seroadaptive behaviours in HIV-positive MSM helped me appreciate the diversity of behaviours and how these are dependent on time and the social context. The mathematical model has offered a theoretical perspective on the mechanisms in which HIV and LGV can interact and has provided me with new ideas on HIV and STI co-infections.

During the PhD I have strengthened my data management and data-analysis skills and in addition learnt the basics of mathematical modelling. An additional outcome of the thesis was to assess the utility of surveillance data in answering research questions. Working with a governmental health institution has been an interesting experience as it has allowed me to observe how surveillance is implemented at a national level and given a better idea of how policy is formulated. My PhD has been a formative period and it has strengthened my analytical and problem solving skills. I have thoroughly enjoyed the opportunity given to me to present and explore ideas, and the PhD process has made me comfortable working independently. After my PhD I would like to use the skills I have developed and gain a more in-depth understanding of the methods that are and could be used in the field of STI epidemiology.

Appendix for Chapter 2

LGV Enhanced Surveillance form from Public Health England (formerly known as Health Protection Agency).

1st page:



LGV Enhanced Surveillance Report Form HPA:LGV6

Health Protection Agency - Centre for Infections

To be completed for confirmed cases of Lymphogranuloma venereum diagnosed in the UK

Reporting clinic | Reported by

Patient clinic number | Date

CONFIDENTIAL

1. **Sex** ☐ Male ☐ Female
2. **Date of birth** (dd/mm/yy)
3. **Postcode - District of residence** (e.g. NW9)
4. **Ethnicity**

White
☐ British ☐ Irish ☐ other

Black or black British
☐ Caribbean ☐ African ☐ other

Asian or Asian British
☐ Indian ☐ Pakistani
☐ Bangladeshi ☐ other

Chinese ☐

Other/mixed - specify

Unknown ☐

5. **Sexuality** ☐ Homosexual ☐ Heterosexual ☐ Bisexual ☐ Unknown
6. **Date of onset of LGV symptoms** (dd/mm/yy)
7. **Date of 1st presentation at the clinic for this episode** (dd/mm/yy)
8. **Reason(s) for attending**
☐ Symptoms ☐ Referral
☐ Contact tracing ☐ Unknown
☐ Routine STI screen

Other - specify

9. **Site of infection (eg. rectal, urethral, lymphnode?**
10. **Proctitis symptoms** - tick all that apply
☐ No proctitis ☐ Rectal discharge
☐ Rectal pain ☐ Bloody stools
☐ Tenesmus ☐ Constipation

Other - specify

11. **Genital symptoms** - tick all that apply
☐ No genital symptoms ☐ Painful lymph nodes
☐ Abcess/Ulcers ☐ Pain on urinating
☐ Swollen lymph nodes ☐ Discharge (penile/vaginal)

Other - specify

12. **Systemic symptoms** - tick all that apply
☐ No systemic symptoms ☐ Muscular pain
☐ Fever ☐ Weight loss
☐ General malaise

Other - specify

13. **Antibiotic treatment**

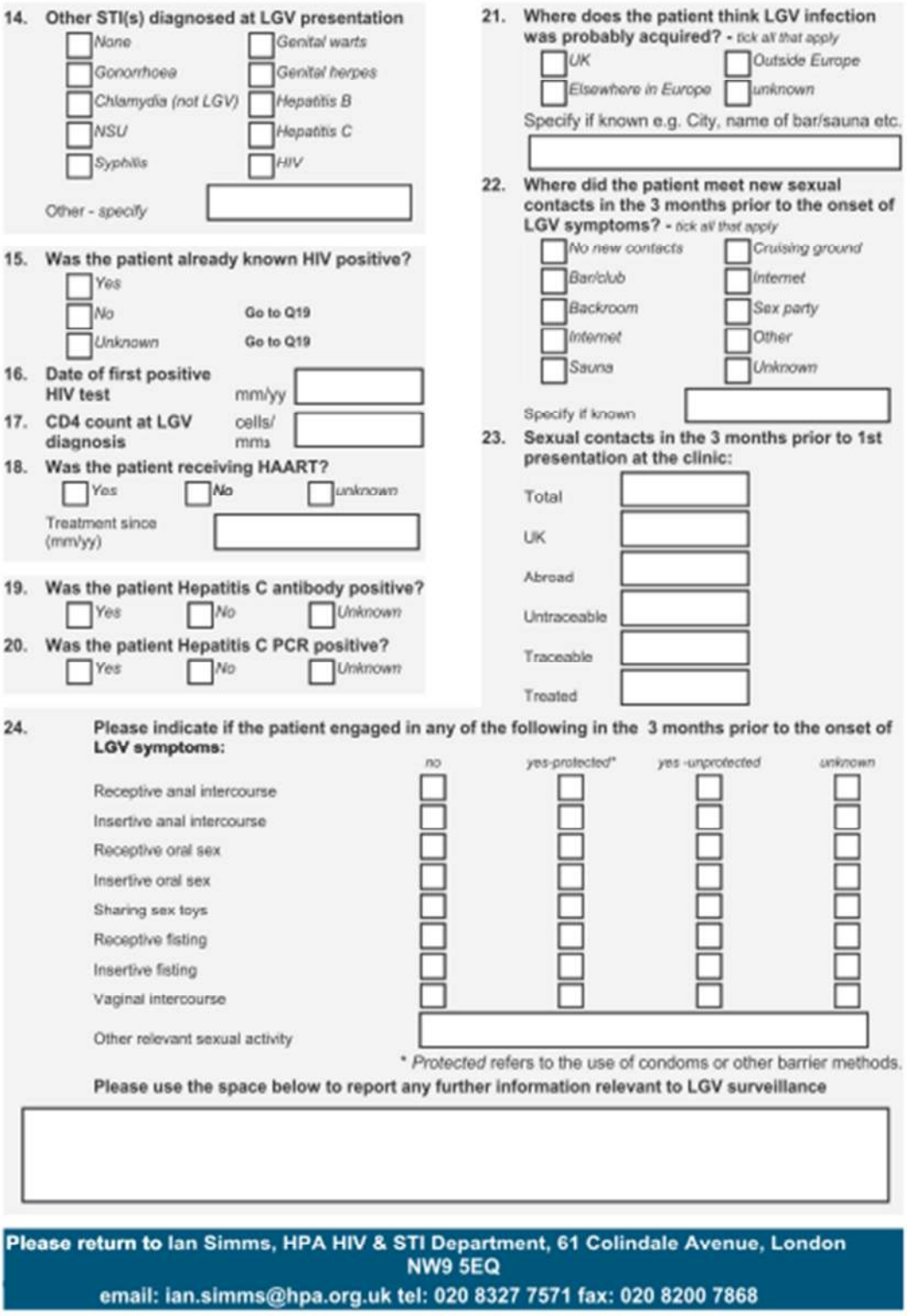
	Dose (mg)	Course (days)
Doxycyline	☐	☐
Azithromycin	☐	☐
Erythromicin	☐	☐

Other - specify

Not treated ☐

CONTINUED OVERLEAF

2nd page:



14. Other STI(s) diagnosed at LGV presentation

- ☐ None
- ☐ Gonorrhoea
- ☐ Chlamydia (not LGV)
- ☐ NSU
- ☐ Syphilis
- ☐ Genital warts
- ☐ Genital herpes
- ☐ Hepatitis B
- ☐ Hepatitis C
- ☐ HIV

Other - specify

15. Was the patient already known HIV positive?

- ☐ Yes
- ☐ No — Go to Q19
- ☐ Unknown — Go to Q19

16. Date of first positive HIV test mm/yy

17. CD4 count at LGV diagnosis cells/mm3

18. Was the patient receiving HAART?

- ☐ Yes
- ☐ No
- ☐ unknown

Treatment since (mm/yy)

19. Was the patient Hepatitis C antibody positive?

- ☐ Yes
- ☐ No
- ☐ Unknown

20. Was the patient Hepatitis C PCR positive?

- ☐ Yes
- ☐ No
- ☐ Unknown

21. Where does the patient think LGV infection was probably acquired? - *tick all that apply*

- ☐ UK
- ☐ Elsewhere in Europe
- ☐ Outside Europe
- ☐ unknown

Specify if known e.g. City, name of bar/sauna etc.

22. Where did the patient meet new sexual contacts in the 3 months prior to the onset of LGV symptoms? - *tick all that apply*

- ☐ No new contacts
- ☐ Bar/club
- ☐ Backroom
- ☐ Internet
- ☐ Sauna
- ☐ Cruising ground
- ☐ Internet
- ☐ Sex party
- ☐ Other
- ☐ Unknown

Specify if known

23. Sexual contacts in the 3 months prior to 1st presentation at the clinic:

Total	
UK	
Abroad	
Untraceable	
Traceable	
Treated	

24. Please indicate if the patient engaged in any of the following in the 3 months prior to the onset of LGV symptoms:

	no	yes-protected*	yes -unprotected	unknown
Receptive anal intercourse	☐	☐	☐	☐
Insertive anal intercourse	☐	☐	☐	☐
Receptive oral sex	☐	☐	☐	☐
Insertive oral sex	☐	☐	☐	☐
Sharing sex toys	☐	☐	☐	☐
Receptive fisting	☐	☐	☐	☐
Insertive fisting	☐	☐	☐	☐
Vaginal intercourse	☐	☐	☐	☐

Other relevant sexual activity

* *Protected* refers to the use of condoms or other barrier methods.

Please use the space below to report any further information relevant to LGV surveillance

Please return to Ian Simms, HPA HIV & STI Department, 61 Colindale Avenue, London NW9 5EQ
email: ian.simms@hpa.org.uk tel: 020 8327 7571 fax: 020 8200 7868

Available via PHE's LGV Enhanced Surveillance website (accessed 24 March, 2013):
http://www.hpa.org.uk/web/HPAweb&Page&HPAwebAutoListName/Page/1191942171559
Direct link to the pdf: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1194947315758

Table S 1. Tabulation of data cleaning process and resulting variables. Item # corresponds to the number on the LGV Enhanced Surveillance Form

Variable	Item #	Type of data	Item response N=1370	Variable created	Comments
City of clinic		Text field	99.0% n=1357	London, outside London	Majority of episodes are diagnosed in London Used clinic name for 13 missing variables (all from Wales).
Clinic name		Text field			
Date reported		Date, string variable			Transformed to time variable
Sex	#1	M, F	99.9% n=1368	Male, not male, missing	
Age (date of birth omitted)	#2	Number	99.1% n=1358	Categorical based on mean (38)	See Figure S 1 for distribution of age
Postcode (omitted)	#3				
Ethnicity	#4	White, Asian, black, Chinese, Egyptian, other, unknown	98.4% n=1348	White, black, Asian, other, unknown	Majority of episodes in white ethnic group
"Sub ethnicity"		Text field			
Sexuality	#5	Homosexual, heterosexual, bisexual, unknown	82.9% n=1136		
Date of onset of symptoms	#6	Date, string variable	99.3% n=1361		Transformed to time variable One episode with year 7200, coded as missing
Date of 1st presentation at the clinic for this episode	#7	Date, string variable			Transformed to time variable, no changes made (implicitly assumed that presentation date was always coded correctly)
		Generated using presentation date (#7)	80.7% n=1106	Presented before 2010, presented on 201	Created to explore the potential impact of the increase in diagnoses during 2010. 9 episodes without presentation date were not included.

		Generated from #6 and #7	72.9% n=999	**Duration of symptoms:** Week or less of symptoms, more than week, unknown	Categorical variable is based on mode of symptoms, see **Error! Reference source not found.** for the distribution of symptom duration. When onset of symptoms occurred after presentation day (n=24 negative value), this was coded as missing
		Generated from entry date (reported by) and #6		**Delay in reporting**	When presentation date coded after form fill date (n=6 with negative delay)) this was coded as missing. See Figure S 2.
Reasons for attending:	#8		97.7% n=1338		
Symptoms		Tick		Yes, no, unknown [40]	
Contact tracing		Tick		Yes, no, unknown [41]	
Routine STI screen		Tick		Yes, no, unknown [42]	There seems to have been some overlap between routine STI screen and routine HIV appointment; in addition some clinicians marked HIV appointment as the only reason for clinic attendance
Referral		Tick		Yes, no, unknown	

[40] Examples from text field: “mentioned symptoms”, “did not feel right”, “symptoms admitted to on questioning after protoscopy finding”, “urethral symptom only”, “unrelated penile rash”, “presented with neurosyphilis …”, “penile ulcer and perianal warts”.
[41] Examples from text field: “also contact of suspected HIV+ male”, “contact of STS” [STS=syphilis], “regular partner both with chlamydia”, “partner has symptoms, having test the same day”, “GC positive and LGV contact” [GC=gonorrhoea].
[42] Examples from text field: “HIV care”, “Routine HIV follow up appointment”, “HIV pos. 1st clinic attendance, offered STI screen”, “Screened at routine HIV visit – asymptomatic at the time”.

Other		Text field			
Site of infection	#9	Text field	41.0% n=562	Categorical variable: 0. Rectal infection (including one episode with rectal and throat infection)[43] 1. Genital infection[44], 2. Both rectal and genital infection[45] 3. Other (including one episode with infection in throat) 4. Unknown/missing	Throat and other sites were combined due to small number of events). The variable was added to the LGV ES form later, and has data mainly from the later years of ES data collection.
Proctitis symptoms	#10		98.4% n=1348		
No proctitis		Tick		Yes, no, unknown	
Rectal pain		Tick		Yes, no, unknown	
Tenesmus		Tick		Yes, no, unknown	
Bloody stools		Tick		Yes, no, unknown	
Constipation		Tick		Yes, no, unknown	
Other		Text field			In the text field 2.1% (29/1370) mentioned rectal bleeding[46] and 5.0% (68/1370) mentioned rectal ulcers

[43] Examples from text field: “Anal", “Perianal", " Perianal Ulcer", "Rectal Abdominal" "Rectal/Anal ulcer" "Rectal & throat".
[44] Examples from text field: “Groin", "Inquinal Lymphnode", “L.N. Aspirate", "Lymphnode biopsy", "Penile Shaft", "Penile, Lymphnodes", "Penis", "Urethra"

[45] Examples from text field: "Rectal & Urethral", "Rectal & lymph nodes", "Rectal & urethral", "Rectal, INGVINAL NODES", "Rectal/ Lmyph node", "Rectal/Lympnode", “Urethral & Rectal"

[46] Examples from text field: “Bloody discharge”, “rectal bleeding”, “PR bleeding” *(PR=peri rectum)*, “bleeding after fisting”, “contact bleeding’.

		Generated from #10.		**Any rectal symptoms:** Yes, no, unknown	Any mention of rectal symptoms in the options above (excluding no proctitis) and the text field Summary variable of rectal symptoms; no proctitis was not used as it is not synonymous with no rectal symptoms and therefore difficult to interpret. Any rectal symptom reduces uncertainty in interpreting certain variables and also includes symptoms mentioned in the text field
Genital symptoms	#11		95.2-95.5% n=1304 to n=1309		
No genital symptoms		Tick		Yes, no, unknown	Where no genital symptoms and a genital symptom were both reported no genital symptom was coded as unknown (n=5)
Abscess/ulcers		Tick		Yes, no, unknown	Based on text fields some rectal ulcers may be reported as genital ulcers (in the absence of rectal ulcer option)
Swollen lymph nodes		Tick		Yes, no, unknown	Few text fields were added which mentioned palpable lymph nodes or “groin gland size”
Painful lymph nodes		Tick		Yes, no, unknown	
Pain on urinating		Tick		Yes, no, unknown	
Discharge (penile/vaginal)		Tick		Yes, no, unknown	
Other		Text field			
		Generated from #11.		**Any genital symptoms:** Yes, no, unknown	None used as reference, and any mention of genital symptoms was coded as yes (including text field)

		Generated from #11 and #12		**Location of symptoms reported:** 0. None 1. Only genital 2. Only rectal 3. Both rectal and genital 4. Unknown/missing	
Systemic symptoms	#12		95.2-95.6% n=1304 to n=1310		
No systemic symptoms		Tick		Yes, no, unknown	Six episodes had no systemic symptoms with another systemic symptom option marked, and these were coded as unknown
Fever		Tick		Yes, no, unknown	
Muscular pain		Tick		Yes, no, unknown	
Weight loss		Tick		Yes, no, unknown	
General malaise		Tick		Yes, no, unknown	
Other		Tick			Fewer text field responses in this section compared to rectal and genital symptoms
		Generated from #12.		**Any systemic symptoms:** Yes, no, unknown	Combined variable of any systemic symptom in the options or in the text field
Antibiotic treatment	#13	Text field			Presented as a text field with treatment regimens; this variable was not modified
Other STIs diagnosed at presentation	#14		94.5-94.7% n=1295 to n=1297		
None		Tick		Yes, no, unknown	Where another STI and “none” were selected, none was recoded as unknown
Gonorrhoea		Tick		Yes, no, unknown	

Chlamydia (not LGV)	Tick	Yes, no, unknown	Significance of the variable questionable as the confirmation of chlamydia subtype only applies to LGV, and it is unlikely most of these would have been confirmed as non-LGV chlamydia.
NSU	Tick	Yes, no, unknown	
Syphilis	Tick	Yes, no, unknown	
Genital warts	Tick	Yes, no, unknown	
Genital herpes	Tick	Yes, no, unknown	Where text field mentioned HSV it was assumed to mean genital herpes (HSV2) and recoded as such
Hepatitis B	Tick	Yes, no, unknown	
Hepatitis C	Tick	Yes, no, unknown	Under questions 19 and 20 there is better diagnostic information of hepatitis C, and this variable was not used for further analysis. There is little concordance between this section and answers from antibody and PCR test results
HIV	Tick	Yes, no, unknown	This field should indicate newly diagnosed HIV infection. When date of first HIV-positive test was compared to presentation date to clinic for these dates, there was some discrepancy between the dates or the first HIV-positive test was often missing. Therefore this cannot be directly viewed as incident HIV diagnoses.
Other	Text field		Few STIs not listed in the options were mentioned in the text field: mycoplasma, scabies, mollusca, entamoeba.
		Any concurrent STI: yes, no, unknown	Including gonorrhoea, syphilis, warts, herpes, hepatitis B, NSU or any of the rarer STIs mentioned in the text field. HIV and hepatitis C were not included as separate variables were created for these. Chlamydia was not included as confirmation of non-LGV chlamydia is unlikely to have happened in most instances (if this

					was the only reported STI the variable was coded as unknown).
Was the patient already known HIV-positive?	#15	Yes, no, unknown	99.7% n=1366	Yes, no, unknown	Interpretation of this question is somewhat ambiguous: known to whom, to the clinician or to the patient?
Date of first positive HIV test	#16	Month, year; string variables	(for HIV variables, denomina tor is not 1370) n=784		Coded from two string variables to a date variable (month, year) One year was 380 (potentially entered the CD4 cell count here) and this was coded as missing. It was suspected this may have been misunderstood in some cases (where the patient was otherwise marked as HIV-negative) as the date of last HIV test. Where no other indication of HIV-positivity existed than the test date, the episode was assumed HIV status unknown.
		Generated using presentation date (#7) and #16, when both available	n=780	**Duration of diagnosed HIV until presentation to the clinic with LGV**	Months between HIV diagnosis and presentation date, see Figure S 5. In two episodes the HIV diagnosis is after the form entry date, and these episodes were coded as missing In seven episodes it seems the first positive HIV test is after presentation to clinic with LGV (in three 1 month, in three in 2-3 months and in one 12 months post presentation). First positive test over month after presentation date was considerate to indicate seroconversion after LGV episode.
CD4 count at LGV diagnosis	#17	Number	n=952		No changes were deemed necessary for this variable. It was assumed a response to the variable was a strong indication of the patient's HIV-status.
Was the patient receiving HAART?	#18	Yes, no, unknown	n=1062	Yes, no, unknown	Categorical variable was created so that tis i nested in the HIV-status variable.
Treatment since		Month, year	Month, n=309 Year, n=387		Not modified and used mainly for HIV status validation

		Generated from #15-18 (and #14. HIV)		**HIV-positivity at LGV diagnosis** Yes, negative/unknown	In the presence of conflicting information the HIV-related questions were evaluated to determine the HIV status, and in uncertain cases, the episodes were deemed to be HIV-negative/unknown. No question confirms the perceived HIV-negative status (last HIV-negative test), therefore everyone who is not HIV-positive based on the responses should be classified as HIV-unknown.
Was the patient hepatitis C antibody positive	#19	Yes, no, unknown	98.8% n=1353	Yes, no, unknown	
Was the patient hepatitis C PCR positive	#20	Yes, no, unknown	96.1% n=1317	Yes, no, unknown	
Where does the patient think LGV infection was probably acquired?	#21	Several text fields indicating UK, Europe or outside	88.0% n=1206	**UK not reported, UK reported, unknown** **Abroad not reported, abroad reported, unknown** **Combined variable:** Only UK reported, only abroad reported, both UK and abroad reported, unknown	In the dataset the locations appear categorised into several text fields with UK sub-locations and locations abroad. Abroad here includes both European and non-European locations.
Where did the patient meet new sexual contacts in the 3 months prior to the onset of LGV symptoms	#22		48.6-48.8% n=666 to n=669		Like with similar questions above, the question was considered answered if any of the boxes (excluding unknown) was answered or the text field filled
No new contacts		Tick		Yes, no, unknown	This was considerate exclusive of other options , and where no new contacts and location were both mentioned this was recoded as unknown
Bar/club		Tick		Yes, no, unknown	Few episodes had specified bars and sex clubs in London under venue text field, and these were recoded under bar/club category.
Backroom		Tick		Yes, no, unknown	

Cruising ground		Tick		Yes, no, unknown	
Internet		Tick		Yes, no, unknown	On the form internet appears twice, but in the dataset there is only one variable.
Sauna		Tick		Yes, no, unknown	Several text fields exist where sauna was reported or a specific sauna was identified.
Sex party		Tick		Yes, no, unknown	
Unknown		Tick		Yes, no, unknown	
Other		Text field			
Specify if known		Text field: venue			Text fields were checked and venue types were corrected if the corresponding option was not marked
				Any location Yes, no, unknown	Correlation between variables.
				Any sex on venue location Yes, no, unknown	Included backroom, cruising ground, sauna and sex party To differentiate the more high-risk venues from the more mainstream venues (internet and bar/club) which may not be a marker of high-risk behaviour in the same way.
Sexual contacts in the past 3 months prior to 1st presentation at the clinic	#23				Nothing was done to these variables, and total number of partners, with the highest coverage, was selected as the variable of choice.
Total		Number	88.4% n=1211	0-3 partners, >3 partners, unknown	Based on the median of 3 (range 0-213). In Figure S 4 we see a rapidly decaying distribution in sexual partner numbers, based on this median was chosen as a cut-off point for categorical variable.
UK		Number	75.0% n=1028		

Abroad		Number	24.8% n=340		
Untraceable		Number	51.3% n=703		
Traceable		Number	53.6% n=735		
Treated		Number	26.3% n=360		
Sexual practices in the 3 months prior to the onset of LGV symptoms	#24				
RAI		No, yes, unknown	99.1% n=1357	**RAI, combined:** 0. Did not report any RAI 1. Reported only RAI (no information on protection) or reported protected RAI 2. Reported unprotected RAI 3. Unknown/missing	Categorisation tried to capture whether RAI was reported, and if it was reported unprotected. Due to the small number of people who reported protected RAI alone, this was combined with reporting RAI without information on protection. Similar approach was taken with the other sexual practices.
RAI, protected		Yes	67.4% n=923)		
RAI, unprotected		Yes	68.3% n=936		
IAI		No, yes, unknown	97.4% n=1335		
IAI, protected		Yes, unknown	26.3% (360		
IAI, unprotected		No, yes, unknown	53.6% n=735		
		Generated from the different IAI variables		**IAI, combined:** 0. Did not report any IAI, 1. Reported only IAI (no information on protection) or reported protected IAI	Those who reported unprotected sex acts may have also reported protected sex in addition but they were categorised according to their riskiest act (for transmission). Similar approach was taken with all combined type of sex variables.

			2. Reported unprotected IAI 3. Unknown/missing	
Receptive oral sex	No, yes, unknown	97.2% n=1332		
Receptive oral sex, protected	Yes	4.4% n=60		
Receptive oral sex, unprotected	Yes, unknown	76.6% n=1050		
	Generated from the different receptive sex variables		**Receptive oral sex, combined:** 0. None reported 1. Reported only receptive oral sex (no information on protection) or reported protected receptive oral sex 2. Reported unprotected receptive oral sex 3. Unknown/missing	
Insertive oral sex	No, yes, unknown	97.2% n=1329		
Insertive oral sex, protected	Yes	95.6% n=57		
Insertive oral sex, unprotected	Yes	74.2% n=1016		
	Generated from the different insertive oral sex variables		**Insertive oral sex, combined:** 0. None reported 1. Reported only receptive oral sex (no information on protection) or reported protected oral sex 2. Reported unprotected receptive oral sex 3. Unknown/missing	
	Generated from insertive and receptive oral sex variables		**Any oral sex:** 0. None reported 1. Reported some (prot or prot unknown)	Combined variables created as there was correlation between variables

			2. Reported one unprotected 3. Reported both unprotected 4. Some or all unknown
Sharing sex toys	No, yes, unknown	94.1% n=1289	
Sharing toys, protected	Yes	0.4% n=5	
Sharing toys, unprotected	Yes	6.3% n=87	
	Generated from the different sharing toys variables		**Sharing toys, combined:** 0. Did not report any sharing of toys 1. Reported only some (no information on protection) or reported protected sharing of toys 2. Reported unprotected sharing (may have also reported protected sharing) 3. Unknown/missing
Receptive fisting	No, yes, unknown	94.2% n=1290	
Fistee, protected	Yes	0.9% n=12	
Fistee, unprotected	Yes, unknown	8.4% n=115	
			Receptive fisting, combined: 0. Did not report any receptive fisting 1. Reported only receptive fisting (no information on protection) or reported it protected 2. Reported unprotected receptive fisting 3. Unknown/missing.

Insertive fisting	No, yes, unknown	94.2% n=1291		
Fister, protected	Yes	1.0% n=14		
Fister, unprotected	Yes	7.5% n=102		
			Insertive fisting, combined: 0. Did not report any insertive fisting 1. Reported only insertive fisting (no information on protection) or reported it protected 2. Reported unprotected insertive fisting 3. Unknown/missing	
	Generated from receptive and insertive fisting		**Any fisting, combined:** 0. None reported 1. Some fisting reported 2. Both reported unprotected 3. Some or all unknown	
Vaginal intercourse	No, yes, yes(unprotected), unknown	93.9% n=1287	Did not report vaginal intercourse, reported vaginal intercourse (including reporting unprotected vaginal intercourse).	Due to a small number of episodes reporting vaginal intercourse a simpler variable was created.
Other relevant sexual activity	Text field	6.4% n=88		Few had commented on other sexual practices. The most common comments related to rimming (in 20 episodes) followed by douching/enema use (15 responses).
Any further information	Text field	28.9% n=396		Handful of responses related to the severity of symptoms and approximately seven responses indicated sex work. Few responses specified sexual assault.

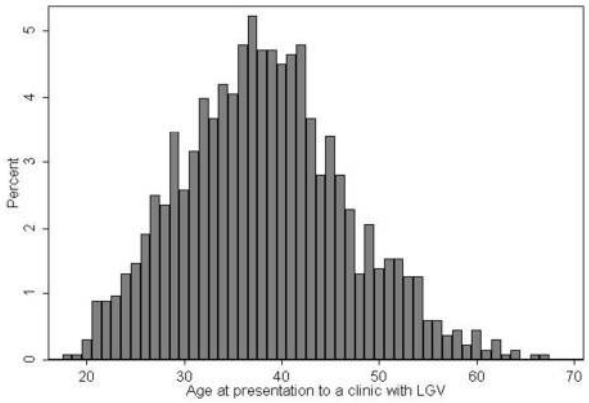


Figure S 1. Distribution of age at the presentation to the clinic.

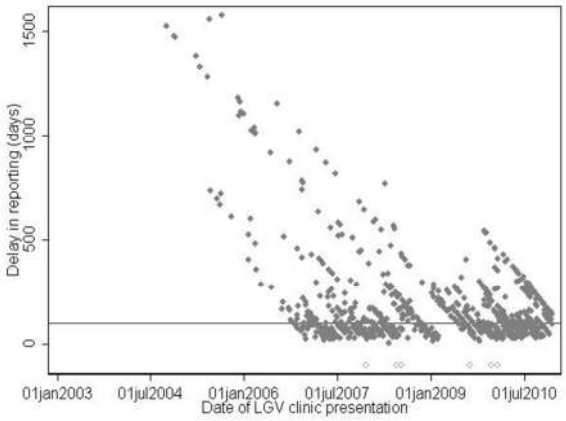


Figure S 2. Scatter plot of delay in reporting in days (y axis) and the presentation to the clinic (x axis), median of delay is 98 days and is presented as black horizontal line on the graph. White circles present episodes where the form filling occurred prior to clinic presentation.

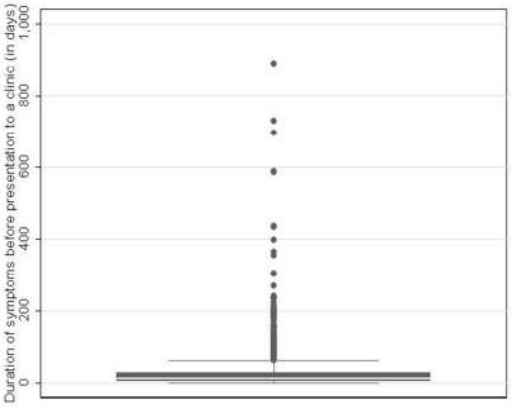


Figure S 3. Boxplot of duration of symptoms in days with a median of 12 (IQR 5-28)

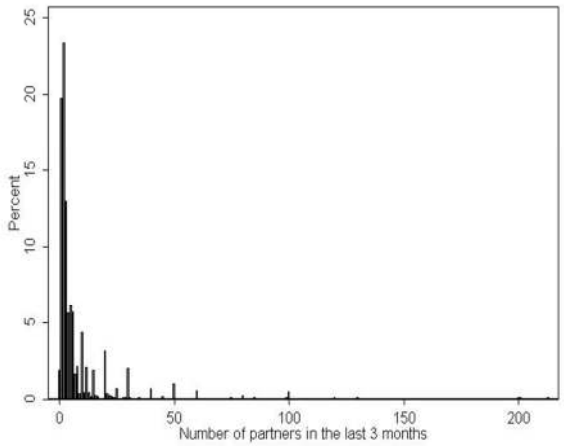


Figure S 4. Number of sexual partners (total) in the past 3 months prior to presentation to clinic. Range is 0-213.

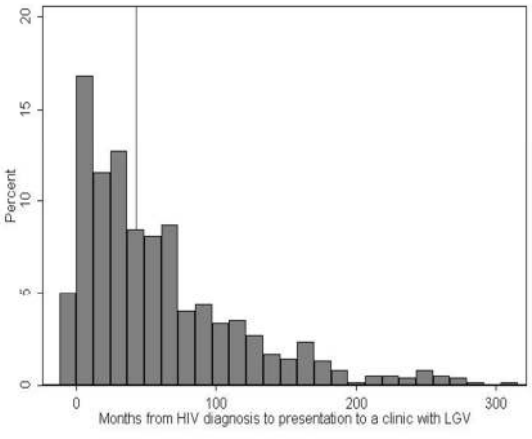


Figure S 5. Duration of "diagnosed" HIV until diagnosed LGV for those with information available. Median duration (43 months) is marked with black vertical line.

Table S 2. Descriptive tabulation of variables in LGV Enhanced Surveillance dataset after data clening. The numbers in brackets refer to the item number in the surveillance form. Tables are presented in several smaller tables numbered S 2.1 to S 2.4.

S 2.1

	n	%	
Seen in a clinic in London			
No	424	30.95	
Yes	946	69.05	
Year of presentation			
Prior to 2010	941	69.14	
During 2010	420	30.86	
	Mean	Median	Range
Age (#2)	38.30	38	18-67
Ethnicity (#4)			
White	1,202	87.74	
Black	63	4.6	
Asian	39	2.85	
Other	43	3.14	
Unknown	23	1.68	
Sexuality (#5)			
Homosexual	1,329	97.01	
Heterosexual	7	0.51	
Bisexual	23	1.68	
Unknown	11	0.8	
	Mean	Median, (mode)	Range
Duration of symptoms (days) (#6)	31.13	12, (7)	0-889
Duration of symptoms			
Week or less	451	32.92	
More than a week	655	47.81	
Unknown	264	19.27	
Reasons for attending (#8)			
Symptoms			
No	174	12.7	
Yes	1,164	84.96	
Unknown	32	2.34	
Contact tracing			
No	1,233	90	
Yes	105	7.66	
Unknown	32	2.34	
Routine STI screen			
No	1,243	90.73	
Yes	95	6.93	
Unknown	32	2.34	
Referral			
No	1,289	94.09	
Yes	49	3.58	
Unknown	32	2.34	

S 2.2

	n	%
Site of infection		
Rectal	536	39.12
Genital	15	1.09
Both or other (throat, n=1)	11	0.80
Missing	808	58.98
Proctitis symptoms		
No proctitis		
No	1,191	86.93
Yes	157	11.46
Unknown	22	1.61
Rectal pain		
No	598	43.65
Yes	750	54.74
Unknown	22	1.61
Tenesmus		
No	1,058	77.23
Yes	290	21.17
Unknown	22	1.61
Rectal discharge		
No	449	32.77
Yes	899	65.62
Unknown	22	1.61
Bloody stools		
No	695	50.73
Yes	653	47.66
Unknown	22	1.61
Constipation		
No	1,099	80.22
Yes	249	18.18
Unknown	22	1.61
Any rectal symptoms		
No	144	10.51
Yes	1,204	87.88
Unknown	22	1.61
Genital symptoms (#11)		
No genital smyptoms		
No	303	22.12
Yes	1,001	73.07
Unknown	66	4.82
Abscess/ulcers		
No	1,217	88.83
Yes	92	6.72
Unknown	61	4.45
Swollen lymph nodes		
No	1,219	88.98
Yes	90	6.57
Unknown	61	4.45
Painful lymph nodes		
No	1,252	91.39
Yes	57	4.16
Unknown	61	4.45
Pain on urinating		
No	1,234	90.07
Yes	75	5.47
Unknown	61	4.45
Discharge (penile/vaginal)		
No	1,260	91.97
Yes	49	3.58
Unknown	61	4.45
Any genital symptoms		
No	1,001	73.07
Yes	307	22.41
Unknown	62	4.53
Systemic symptoms reported (#12)		
No systemic symptoms		
No	351	25.62
Yes	953	69.56
Unknown	66	4.82
Fever		
No	1,178	85.99
Yes	132	9.64
Unknown	60	4.38
Muscular pain		
No	1,253	91.46
Yes	57	4.16
Unknown	60	4.38
Weight loss		
No	1,222	89.2
Yes	88	6.42
Unknown	60	4.38
General malaise		
No	1,100	80.29
Yes	210	15.33
Unknown	60	4.38
Any systemic symptoms		
No	953	69.56
Yes	357	26.06
Unknown	60	4.38

S 2.3

	n	%
Concurrent STIs reported (#14)		
No other STIs		
No	525	38.32
Yes	770	56.2
Unknown	75	5.47
Gonorrhoea		
No	1,061	77.45
Yes	236	17.23
Unknown	73	5.33
Chlamydia		
No	1,189	86.79
Yes	107	7.81
Unknown	74	5.4
NSU		
No	1,263	92.19
Yes	34	2.48
Unknown	73	5.33
Syphilis		
No	1,215	88.69
Yes	82	5.99
Unknown	73	5.33
Warts		
No	1,248	91.09
Yes	49	3.58
Unknown	73	5.33
Herpes		
No	1,260	91.97
Yes	37	2.7
Unknown	73	5.33
Hepatitis B		
No	1,296	94.6
Yes	1	0.07
Unknown	73	5.33
Any other STI		
No	820	59.85
Yes	399	29.12
Unknown	151	11.02
Hepatitis C (from #14 other STIs)		
No	1,269	92.63
Yes	28	2.04
Unknown	73	5.33
Hepatitis C (PCR) (#19)		
No	477	34.82
Yes	145	10.58
Unknown	748	54.6
Hepatitis C (Ab) (#20)		
No	880	64.23
Yes	219	15.99
Unknown	271	19.78
HIV-positivity (#14-18)		
No	220	16.06
Yes	1,109	80.95
Unknown	41	2.99

	N	Mean	Median	Range
CD4 cell count (#16)	951	532.54	500	23-1700

HAART (#18)		
No	424	38.23
Yes	596	53.74
Unknown	89	8.03

	n	%
Acquistion country UK (#21)		
No	99	7.23
Yes	1,107	80.8
Unknown	164	11.98
Acquisition country aborad (#21)		
No	1,044	76.2
Yes	162	11.82
Unknown	164	11.98
Acquisition country (#21)		
UK	1,044	76.2
Abroad	99	7.23
Either	63	4.6
Unknown	164	11.97
Locations for meeting new partners (#22)		
No new partners		
No	576	42.04
Yes	90	6.57
Unknown	704	51.39
Met new partners in bar/club		
No	400	29.2
Yes	269	19.64
Unknown	701	51.17
Met new partners in bakcroom		
No	606	44.23
Yes	63	4.6
Unknown	701	51.17
Met new partners on internet		
No	467	34.09
Yes	202	14.74
Unknown	701	51.17
Met new partners in sauna		
No	483	35.26
Yes	186	13.58
Unknown	701	51.17
Met new partners in cruising ground		
No	621	45.33
Yes	48	3.5
Unknown	701	51.17
Met new partners at sex party		
No	554	40.44
Yes	115	8.39
Unknown	701	51.17
Met partners in any location		
No	90	6.57
Yes	579	42.26
Unknown	701	51.17
Met partners in sex on venue location		
No	347	25.33
Yes	322	23.5
Unknown	701	51.17

Number of partners in the prior three months (#23)

	N	Mean	Median	Range
Total	1211	7.88	3	0-213
Partners in UK	1028	6.80	2	0-201
Partners abroad	340	2.99	1	0-213
Untracable	703	8.07	3	0-200
Tracable	735	2.23	1	0-34
Treated	360	1.13	1	0-23

S 2.4

	n	%
Sexual practices in the past 3 months (#24)		
Receptive anal intercourse		
None reported	82	5.99
Reported protected/or unk	241	17.59
Unprotected	936	68.32
Unknown	111	8.1
Insertive anal intercourse		
None reported	130	9.49
Reported protected/or unk	220	16.06
Unprotected	733	53.5
Unknown	287	20.95
Receptive oral sex		
None reported	112	8.18
Reported protected/or unk	38	2.77
Unprotected	1,045	76.28
Unknown	175	12.77
Insertive oral sex		
None reported	134	9.78
Reported protected/or unk	33	2.41
Unprotected	1,013	73.94
Unknown	190	13.87
Any oral sex		
None reported	104	7.59
Reported some	29	2.12
Reported one unprotected	40	2.92
Reported both unrprotected	993	72.48
Some or all unknown	204	14.89
Receptive fisting		
None reported	571	41.68
Reported protected/or unk	34	2.48
Unprotected	114	8.32
Unknown	651	47.52
Insertive fisting		
None reported	589	42.99
Reported protected/or unk	15	1.09
Unprotected	102	7.45
Unknown	664	48.47
Any fisting		
No fisting reported	547	39.93
Some fisting reported	62	4.53
Both reported, unprotected	79	5.77
Some unknown	682	49.78
Sharing sex toys		
No	531	38.76
Any (prot or unpr)	96	7.01
Unknown	743	54.23
Vaginal intercourse		
No	1,015	74.09
Yes (unprot or not)	13	0.95
Unknown	342	24.96

Table S 3. Variables which were collapsed into combined variables.

	HIV+ (n=1087)		HIV-/unknown (n=255)		Univariate logistic regression (GEE)			
	n	%	n	%	OR	1.0	CI	P-value
Combined variable: any rectal symptoms								
No	103	**9.5**	35	**13.7**	1.0			
Yes	974	**89.6**	209	**82.0**	1.6	1.1	2.4	0.026
Unknown	10	**0.9**	11	**4.3**	0.3	0.1	0.8	0.019
Combined variable: any genital symptoms								
No	823	**75.7**	162	**63.5**	1.0			
Yes	215	**19.8**	83	**32.6**	0.5	0.4	0.7	0.000
Unknown	49	**4.5**	10	**3.9**	1.0	0.5	2.0	0.957
Systemic symptoms reported:								
No systemic symptoms								
No	286	**26.3**	57	**22.4**	1.0			
Yes	752	**69.2**	186	**72.9**	0.8	0.9	1.1	0.191
Unknown	49	**4.5**	12	**4.7**	0.9	0.4	1.7	0.649
Fever								
No	929	**85.5**	225	**88.2**	1.0			
Yes	113	**10.4**	18	**7.1**	1.5	0.9	2.5	0.116
Unknown	45	**4.1**	12	**4.7**	0.9	0.5	1.7	0.774
Muscular pain								
No	993	**91.4**	236	**92.6**	1.0			
Yes	49	**4.5**	7	**2.8**	1.7	0.8	3.6	0.211
Unknown	45	**4.1**	12	**4.7**	0.9	0.5	1.7	0.734
Weight loss								
No	973	**89.5**	228	**89.4**	1.0			
Yes	69	**6.4**	15	**5.9**	1.1	0.6	1.9	0.815
Unknown	45	**4.1**	12	**4.7**	0.9	0.5	1.7	0.702
Malaise								
No	869	**79.9**	213	**83.5**	1.0			
Yes	173	**15.9**	30	**11.8**	1.4	0.9	2.1	0.106
Unknown	45	**4.1**	12	**4.7**	0.9	0.5	1.8	0.804
Concurrent STIs reported								
No other STIs								
No	415	**38.2**	94	**36.9**	1.0			
Yes	618	**56.9**	140	**54.9**	1.0	0.8	1.3	0.994
Unknown	54	**5.0**	21	**8.2**	0.6	0.3	1.0	0.058
Gonorrhoea								
No	842	**77.5**	197	**77.3**	1.0			
Yes	192	**17.7**	38	**14.9**	1.2	0.8	1.7	0.386
Unknown	53	**4.9**	20	**7.8**	0.6	0.4	1.1	0.091
Chlamydia								
No	955	**87.9**	209	**82.0**	1.0			
Yes	78	**7.2**	26	**10.2**	0.7	0.4	1.1	0.098
Unknown	54	**5.0**	20	**7.8**	0.6	0.4	1.0	0.061
NSU								
No	1,008	**92.7**	227	**89.0**	1.0			
Yes	26	**2.4**	8	**3.1**	0.7	0.3	1.7	0.470
Unknown	53	**4.9**	20	**7.8**	0.6	0.4	1.0	0.065
Syphilis								
No	968	**89.1**	222	**87.1**	1.0			
Yes	66	**6.1**	13	**5.1**	1.1	0.6	2.1	0.658
Unknown	53	**4.9**	20	**7.8**	0.6	0.4	1.0	0.075
Warts								
No	995	**91.5**	226	**88.6**	1.0			
Yes	39	**3.6**	9	**3.5**	1.0	0.5	2.1	0.982
Unknown	53	**4.9**	20	**7.8**	0.6	0.4	1.0	0.070
Herpes								
No	1,007	**92.6**	227	**89.0**	1.0			
Yes	27	**2.5**	8	**3.1**	0.8	0.3	1.7	0.514
Unknown	53	**4.9**	20	**7.8**	0.6	0.4	1.0	0.065

	HIV+ (n=1087)		HIV-/unknown (n=255)		Univariate logistic regression (GEE)			
	n	%	n	%	OR	1.0	CI	P-value
Locations for meeting new partners								
No new partners								
No	461	**42.4**	100	**39.2**	1.0			
Yes	76	**7.0**	13	**5.1**	1.3	0.7	2.3	0.469
Unknown	550	**50.6**	142	**55.7**	0.8	0.6	1.1	0.232
Met new partners in bar/club								
No	334	**30.7**	56	**22.0**	1.0			
Yes	206	**19.0**	57	**22.4**	0.6	0.4	0.9	0.017
Unknown	547	**50.3**	142	**55.7**	0.7	0.5	0.9	0.012
Met new partners in bakcroom								
No	487	**44.8**	105	**41.2**	1.0			
Yes	53	**4.9**	8	**3.1**	1.4	0.7	3.0	0.357
Unknown	547	**50.3**	142	**55.7**	0.8	0.6	1.1	0.199
Met new partners on internet								
No	371	**34.1**	83	**32.6**	1.0			
Yes	169	**15.6**	30	**11.8**	1.2	0.8	1.9	0.320
Unknown	547	**50.3**	142	**55.7**	0.9	0.6	1.2	0.334
Met new partners in sauna								
No	387	**35.6**	86	**33.7**	1.0			
Yes	153	**14.1**	27	**10.6**	1.2	0.8	2.0	0.361
Unknown	547	**50.3**	142	**55.7**	0.9	0.6	1.1	0.302
Met new partners in cruising ground								
No	500	**46.0**	107	**42.0**	1.0			
Yes	40	**3.7**	6	**2.4**	1.4	0.6	3.3	0.422
Unknown	547	**50.3**	142	**55.7**	0.8	0.6	1.1	0.180
Met new partners at sex party								
No	445	**40.9**	98	**38.4**	1.0			
Yes	95	**8.7**	15	**5.9**	1.4	0.8	2.5	0.234
Unknown	547	**50.3**	142	**55.7**	0.9	0.6	1.1	0.274
Receptive oral sex								
None reported	77	**7.1**	31	**12.2**	1.0			
Reported protected/or prot.unk.	29	**2.7**	9	**3.5**	1.3	0.6	3.0	0.557
Unprotected	850	**78.2**	176	**69.0**	1.9	1.2	3.0	0.004
Unknown	131	**12.1**	39	**15.3**	1.3	0.8	2.3	0.290
Insertive oral sex								
None reported	96	**8.8**	33	**12.9**	1.0			
Reported protected/or prot.unk.	26	**2.4**	7	**2.8**	1.3	0.5	3.2	0.598
Unprotected	823	**75.7**	174	**68.2**	1.6	1.0	2.4	0.029
Unknown	142	**13.1**	41	**16.1**	1.2	0.7	2.0	0.530
Receptive fisting								
None reported	439	**40.4**	118	**46.3**	1.0			
Reported protected/or prot.unk.	24	**2.2**	9	**3.5**	0.7	0.3	1.6	0.399
Unprotected	100	**9.2**	10	**3.9**	2.6	1.3	4.9	0.004
Unknown	524	**48.2**	118	**46.3**	1.2	0.9	1.6	0.208
Insertive fisting								
None reported	455	**41.9**	119	**46.7**	1.0			
Reported protected/or prot.unk.	12	**1.1**	3	**1.2**	1.1	0.3	3.9	0.918
Unprotected	92	**8.5**	8	**3.1**	2.9	1.4	5.9	0.003
Unknown	528	**48.6**	125	**49.0**	1.1	0.8	1.5	0.461

Table S 4. Reported symptoms and site of infection cross-tabulated and stratified by HIV status

HIV-negative/unknown (total)

		Reported symptoms					
Site of infection		None	Only genital	Only rectal	Both	One or both unknown	Total
Rectal	n	5	2	42	16	1	66
	%	**7.6**	**3.0**	**63.6**	**24.2**	**1.5**	**100.0**
Genital	n	0	4	0	1	0	5
	%	**0.0**	**80.0**	**0.0**	**20.0**	**0.0**	**100.0**
Both/throat	n	1	0	0	2	0	3
	%	**33.3**	**0.0**	**0.0**	**66.7**	**0.0**	**100.0**
Unknown	n	5	18	105	35	18	181
	%	**2.8**	**9.9**	**58.0**	**19.3**	**9.9**	**100.0**
Total	n	11	24	147	54	19	255
	%	**4.3**	**9.4**	**57.7**	**21.2**	**7.5**	**100.0**

HIV-positive (total)

		Reported symptoms					
Site of infection		None	Only genital	Only rectal	Both	One or both unknown	Total
Rectal	n	25	12	319	76	23	455
	%	**5.5**	**2.6**	**70.1**	**16.7**	**5.1**	**100.0**
Genital	n	1	8	0	0	1	10
	%	**10.0**	**80.0**	**0.0**	**0.0**	**10.0**	**100.0**
Both/throat	n	0	1	2	5	0	8
	%	**0.0**	**12.5**	**25.0**	**62.5**	**0.0**	**100.0**
Unknown	n	34	20	437	91	32	614
	%	**5.5**	**3.3**	**71.2**	**14.8**	**5.2**	**100.0**
Total	n	60	41	758	172	56	1,087
	%	**5.5**	**3.8**	**69.7**	**15.8**	**5.2**	**100.0**

Table S 5. Comparison of standard logistic regression and logistic regression with GEE: clinical variables

	HIV positive n=1087		HIV unknown n=255		Univariate logistic regression (GEE)				Univariate logistic regression (standard logistic)			
	n	%	n	%	OR	1.0	CI	P-value	OR	1.0	CI	P-value
Episode number					not included							
1st	1023	**94.1**	254	**99.6**								
2nd	58	**5.3**	1	**0.4**								
3rd	6	**0.6**	0	**0.0**								
Seen in a clinic in London												
No	316	**29.1**	100	**39.2**	1.0							
Yes	771	**70.9**	155	**60.8**	1.5	1.2	2.1	0.003	1.6	1.2	2.1	0.002
Duration of symptoms												
Week or less	377	**34.7**	65	**25.5**	1.0				1.0			
More than a week	500	**46.0**	141	**55.3**	0.6	0.5	0.8	0.002	0.6	0.4	0.8	0.003
Unknown	210	**19.3**	49	**19.2**	0.7	0.5	1.1	0.146	0.7	0.5	1.1	0.145
Presented to the clinic because of												
Reasons for attending the clinic												
Symptoms												
No	139	**12.8**	31	**12.2**	1.0				1.0			
Yes	925	**85.1**	218	**85.5**	0.9	0.6	1.4	0.783	0.9	0.6	1.4	0.795
Unknown	23	**2.1**	6	**2.4**	0.8	0.3	2.1	0.710	0.9	0.3	2.3	0.754
Contact tracing												
No	980	**90.2**	229	**89.8**	1.0				1.0			
Yes	84	**7.7**	20	**7.8**	1.0	0.6	1.6	0.927	1.0	0.6	1.6	0.942
Unknown	23	**2.1**	6	**2.4**	0.9	0.4	2.1	0.769	0.9	0.4	2.2	0.813
Routine STI screen												
No	988	**90.9**	232	**91.0**	1.0				1.0			
Yes	76	**7.0**	17	**6.7**	1.1	0.6	1.8	0.849	1.0	0.6	1.8	0.861
Unknown	23	**2.1**	6	**2.4**	0.9	0.4	2.1	0.779	0.9	0.4	2.2	0.821
Referral												
No	1,033	**95.0**	233	**91.4**	1.0				1.0			
Yes	31	**2.9**	16	**6.3**	0.4	0.2	0.8	0.011	0.4	0.2	0.8	0.009
Unknown	23	**2.1**	6	**2.4**	0.8	0.4	2.0	0.713	0.9	0.3	2.1	0.754
Location of symptoms reported												
None	60	**5.5**	11	**4.3**	1.0				1.0			
Only Genital	41	**3.8**	24	**9.4**	0.3	0.1	0.7	0.005	0.3	0.1	0.7	0.005
Only Rectal	758	**69.7**	147	**57.7**	1.0	0.5	1.8	0.880	0.9	0.5	1.8	0.869
Both	172	**15.8**	54	**21.2**	0.6	0.3	1.2	0.140	0.6	0.3	1.2	0.139
Unknown	56	**5.2**	19	**7.5**	0.6	0.3	1.3	0.167	0.5	0.2	1.2	0.145
Site of infection												
Rectal	454	**41.8**	63	**29.3**	1.0				1.0			
Genital	10	**0.9**	5	**2.3**	0.3	0.1	0.9	0.032	0.3	0.1	0.9	0.028
Both or other (throat, n=1)	8	**0.7**	3	**1.2**	0.4	0.1	1.5	0.182	0.4	0.1	1.5	0.168
Missing	614	**56.5**	144	**67.0**	0.5	0.4	0.7	<0.001	0.5	0.4	0.7	<0.001
Any systemic symptom												
No	750	**69.0**	186	**72.9**	1.0				1.0			
Yes	292	**26.9**	57	**22.4**	1.3	0.9	1.7	0.150	1.3	0.9	1.8	0.150
Unknown	45	**4.1**	12	**4.7**	0.9	0.5	1.8	0.833	0.9	0.5	1.8	0.829
Any other STI												
No	664	**61.1**	143	**56.1**	1.0				1.0			
Yes	315	**29.0**	72	**28.2**	0.9	0.7	1.3	0.718	0.9	0.7	1.3	0.710
Unknown	108	**9.9**	40	**15.7**	0.6	0.4	0.9	0.010	0.6	0.4	0.9	0.009
Hepatitis C (PCR)												
No	410	**37.7**	57	**22.4**	1.0				1.0			
Yes	138	**12.7**	3	**1.2**	5.9	2.0	17.5	0.001	6.4	2.0	20.7	0.002
Unknown	539	**49.6**	195	**76.5**	0.4	0.3	0.5	<0.001	0.4	0.3	0.5	<0.001

Table S 6 Comparison of standard logistic regression and logistic regression with GEE: behavioural variables

	HIV+ (n=1087)		HIV-/unknown (n=255)		Univariate logistic regression (GEE)				Univariate logistic regression (standard logistic)			
	n	%	n	%	OR	1.0	CI	P-value	OR	1.0	CI	P-value
Age												
mean (sd)	38.60 (8.08)		37.00 (9.92)		1.0	1.0	1.0	0.023	1.0	1.0	1.0	0.008
Ethnicity												
White	953	**87.7**	229	**89.8**	1.0				1.0			
Black	51	**4.7**	9	**3.5**	1.3	0.7	2.8	0.426	1.4	0.7	2.8	0.403
Asian	30	**2.8**	7	**2.8**	1.0	0.4	2.3	0.969	1.0	0.4	2.4	0.945
Other	36	**3.3**	7	**2.8**	1.2	0.5	2.7	0.650	1.2	0.5	2.8	0.614
Unknown	17	**1.6**	3	**1.2**	1.3	0.4	4.4	0.625	1.4	0.4	4.7	0.624
Sexuality												
Homosexual	1,076	**99.0**	243	**95.3**	1.0				1.0			
Bisexual	11	**1.0**	12	**4.7**	0.2	0.1	0.5	<0.001	0.2	0.1	0.5	<0.001
Acquisition country												
UK	828	**76.2**	198	**77.7**	1.0				1.0			
Abroad	81	**7.5**	16	**6.3**	1.2	0.7	2.1	0.476	1.2	0.7	2.1	0.502
Either	47	**4.3**	13	**5.1**	0.9	0.5	1.6	0.651	0.9	0.5	1.6	0.652
Unknown/ missing	131	**12.1**	28	**11.0**	1.1	0.7	1.7	0.604	1.1	0.7	1.7	0.614
Met partners in any of the locations												
No	76	**7.0**	13	**5.1**	1.0				1.0			
Yes	464	**42.7**	100	**39.2**	0.8	0.4	1.5	0.478	0.8	0.4	1.5	0.470
Unknown	547	**50.3**	142	**55.7**	0.7	0.4	1.2	0.188	0.7	0.4	1.2	0.185
Number of contacts												
mean (sd)	8.27 (16.66)		6.56 (17.18)		1.0	1.0	1.0	0.354	1.0	1.0	1.0	0.167
median (range)	3 (0-201)		3 (0-213)									
Receptive anal intercourse												
None reported	51	**4.7**	24	**9.4**	1.0				1.0			
Reported protected/or unk	161	**14.8**	79	**31.0**	0.9	0.5	1.6	0.840	1.0	0.6	1.7	0.883
Unprotected	791	**72.8**	127	**49.8**	2.9	1.7	4.8	<0.001	2.9	1.7	4.9	<0.001
Unknown	84	**7.7**	25	**9.8**	1.6	0.8	3.0	0.180	1.6	0.8	3.1	0.173
Insertive anal intercourse												
None reported	100	**9.2**	25	**9.8**	1.0				1.0			
Reported protected/or unk	154	**14.2**	65	**25.5**	0.6	0.4	1.0	0.056	0.6	0.4	1.0	0.051
Unprotected	615	**56.6**	106	**41.6**	1.5	0.9	2.3	0.117	1.5	0.9	2.4	0.132
Unknown	218	**20.1**	59	**23.1**	0.9	0.6	1.6	0.811	0.9	0.5	1.6	0.767
Any oral sex												
None reported	72	**6.6**	28	**11.0**	1.0				1.0			
Reported some	23	**2.1**	6	**2.4**	1.5	0.6	3.9	0.439	1.5	0.5	4.0	0.433
Reported one unprotected	30	**2.8**	10	**3.9**	1.2	0.5	2.7	0.728	1.2	0.5	2.7	0.719
Reported both unrprotected	812	**74.7**	165	**64.7**	1.9	1.2	3.0	0.008	1.9	1.2	3.1	0.007
Some or all unknown	150	**13.8**	46	**18.0**	1.3	0.7	2.2	0.414	1.3	0.7	2.2	0.395
Any fisting												
No fisting reported	419	**38.6**	114	**44.7**	1.0				1.0			
Some fisting reported	52	**4.8**	9	**3.5**	1.6	0.8	3.2	0.224	1.6	0.8	3.3	0.229
Both reported, unprotected	71	**6.5**	6	**2.4**	3.1	1.4	6.9	0.007	3.2	1.4	7.6	0.008
Some unknown	545	**50.1**	126	**49.4**	1.2	0.9	1.6	0.245	1.2	0.9	1.6	0.260
Sharing sex toys												
No	412	**37.9**	108	**42.4**	1.0				1.0			
Any (prot or unpr)	77	**7.1**	17	**6.7**	1.2	0.7	2.0	0.581	1.2	0.7	2.1	0.552
Unknown/ missing	598	**55.0**	130	**51.0**	1.2	0.9	1.6	0.193	1.2	0.9	1.6	0.197
Vaginal intercourse												
No	798	**73.4**	201	**78.8**	1.0				1.0			
Yes (unprot or not)	4	**0.4**	2	**0.8**	0.5	0.1	2.6	0.415	0.5	0.1	2.8	0.430
Unknown/missing	285	**26.2**	52	**20.4**	1.4	1.0	1.9	0.051	1.4	1.0	1.9	0.058

Table S 7. Covariance matrices of the multivariable models. Section A presents correlation matrix for variables included in the clinical multivariate model, section B has variables included in the behavioural multivariate model and section C has the variables included in the final multivariate model combining behavioural and clinical characteristics.

A	HIV status	Clinic in London	Duration of symptoms	Referral	Location of symptoms reported	Infection site	Any systematic symptom	Any other STI	Form delay (days)	Episode in 2010
HIV status	1.00									
Clinic in London	0.08	1.00								
Duration of symptoms	-0.08	-0.07	1.00							
Referral	-0.07	-0.01	0.11	1.00						
Location of symptoms reported	-0.01	0.05	-0.17	0.06	1.00					
Infection site	-0.12	0.00	0.06	0.13	0.02	1.00				
Any systematic symptom	0.03	0.02	-0.03	0.08	0.25	0.07	1.00			
Any other STI	-0.07	-0.01	0.04	0.04	0.08	0.08	0.00	1.00		
Form delay (days)	-0.03	0.06	0.01	0.00	0.00	0.12	0.01	-0.05	1.00	
Episode in 2010	0.07	-0.03	-0.02	-0.06	-0.06	-0.61	-0.06	-0.05	-0.18	1.00

B	HIV status	Age (years)	Sexuality	Probable acquisition country	Any sex on venue location	Number of contacts	RAI	IAI	Any oral sex	Any fisting	Form delay (days)	Episode in 2010
HIV status	1.00											
Age (years)	0.07	1.00										
Sexuality	-0.14	0.02	1.00									
Probable acquisition country	-0.01	-0.02	-0.04	1.00								
Any sex on venue location	-0.06	-0.02	0.03	0.14	1.00							
Number of contacts	0.07	0.03	-0.03	0.05	-0.02	1.00						
RAI	0.17	0.01	-0.09	0.09	0.06	0.07	1.00					
IAI	0.07	-0.06	-0.05	0.15	0.10	0.00	0.45	1.00				
Any oral sex	0.04	0.02	-0.08	0.12	0.06	0.03	0.26	0.21	1.00			
Any fisting	0.02	0.03	-0.03	0.09	0.20	-0.03	0.11	0.20	0.29	1.00		
Form delay (days)	-0.03	0.01	0.02	0.00	0.06	0.05	0.03	0.04	0.01	0.09	1.00	
Episode in 2010	0.08	0.02	-0.02	0.00	0.01	-0.08	-0.02	-0.10	0.06	0.10	-0.18	1.00

C	HIV status	London	Duration of symptoms	Referral	Location of symptoms reported	Infection site	Any systematic symptom	Any other STI	Age	Sexuality	Probable acquisition country	Number of contacts	RAI	IAI	Any oral sex	Any fisting	Form delay (days)	Episode in 2010
HIV status	1.00																	
London	0.08	1.00																
Duration of symptoms	-0.09	-0.08	1.00															
Referral	-0.07	-0.01	0.10	1.00														
Location of symptoms reported	0.00	0.07	-0.18	0.03	1.00													
Infection site	-0.12	0.01	0.04	0.12	0.03	1.00												
Any systematic symptom	0.05	0.03	-0.03	0.04	0.24	0.06	1.00											
Any other STI	-0.05	0.00	0.04	0.02	0.07	0.08	0.00	1.00										
Age	0.07	-0.05	-0.07	0.00	0.04	-0.05	0.01	-0.08	1.00									
Sexuality	-0.14	-0.04	0.02	0.06	0.04	0.03	0.00	-0.02	0.02	1.00								
Probable acquisition country	-0.01	-0.03	0.04	0.05	0.03	0.02	0.04	0.03	-0.02	-0.04	1.00							
Number of contacts	0.07	0.11	-0.06	-0.05	-0.02	0.06	0.03	0.06	0.03	-0.03	0.05	1.00						
RAI	0.17	0.13	-0.01	0.07	0.11	0.00	0.06	-0.02	0.01	-0.09	0.09	0.07	1.00					
IAI	0.07	0.10	0.03	0.03	0.04	0.05	0.03	-0.02	-0.06	-0.05	0.15	0.00	0.45	1.00				
Any oral sex	0.04	0.01	-0.03	0.03	0.00	-0.10	-0.05	0.03	0.02	-0.08	0.12	0.03	0.26	0.21	1.00			
Any fisting	0.02	-0.05	0.05	-0.04	-0.05	-0.20	-0.06	-0.06	0.03	-0.03	0.09	-0.03	0.11	0.20	0.29	1.00		
Form delay (days)	-0.03	0.08	0.02	0.02	0.00	0.13	0.02	-0.06	0.01	0.02	0.00	0.05	0.03	0.04	0.01	0.09	1.00	
Episode in 2010	0.08	-0.04	-0.01	-0.05	-0.07	-0.62	-0.06	-0.05	0.02	-0.02	0.00	-0.08	-0.02	-0.10	0.06	0.10	-0.18	1.00

Appendix for chapter 3

Description of dataset structure regarding the repeat infections and missing data

In this analysis I included 1342 episodes in 1281 MSM that were included in the analysis in chapter 3. There were 8 repeat episodes that had been excluded since they occurred under 3 months after the previous episode (with presentation day to the clinic as the guideline variable) and they were also excluded from this analysis (and these individuals were not considered repeaters unless there was an additional episode).

Table S 1. Description of missing data on repeat episodes in the LGV Enhanced Surveillance.

Number of episodes in the dataset	Number of individuals
One episode	
1st, 2nd missing	5
2nd, 1st missing	3
3rd, 1st and 2nd missing	1
Two episodes	
1st and 3rd, 2nd missing	1
2nd and 3rd, 1st missing	1
Three episodes	
Episodes 1-3, 4th missing	1
Total	12

Table S 2. Distribution of episodes

Episodes	n	%
Non-repeater	1,215	90.54
Repeater, 1st episode	62	4.62
Repeater, 2nd episode	59	4.40
Repeater, 3rd episode	6	0.45
Total	1,342	100.00

Of those that were known repeaters, the following information had LGV Enhanced Surveillance data attached: Altogether 66 individuals who were known to be re-infected had at least one episode in the dataset in the following way: for 9 individuals there is one episode, for 53 individuals there are 2 episodes and for 4 individuals there are 3 episodes (altogether 127 episodes). For many individuals with a re-infection there is

partial information in the dataset which can be observed in the datasheet given by HPA with lab id and enhanced surveillance id which was used to link episodes known to belong to the same individual as described in chapter 3. The partial information is described in Table S 8, where altogether 12 repeaters have only some of their (known) episodes in the enhanced surveillance dataset (varying from 1 to 3 episodes) making 18.2% (12/66) of repeaters having an episode missing or 9.4% episodes (12/127) without any information. The datasheet used for linkage also revealed 2 individuals with repeat infections (2 and 3) who had no episodes with enhanced surveillance dataset. Distribution of episodes in the dataset by their estimated episode number is presented in Table S 9.

It is also important to consider that the assigned episode number is based on estimate from the presentation day to the clinic and laboratory and enhanced surveillance ID (both are running numbers). One individual was identified where a text field indicated an additional episode between the 1st and 2nd episodes observed in the data, but there was no official linkage based on the laboratory and ES data and the datasheet for linkage was used as the main source for determining episode's status as part of repeat infection series or not.

Table S 3. All variables analysed. Comparison are between repeaters' 1st episode and non-repeaters

	Non-repeaters		Repeaters' 1st episode		Repeaters' 2nd episode		Repeaters' 3rd episode		Univariate logistic regression comparing repeaters' 1st episode to non-repeaters				Fisher's exact test 2-sided p-value for 2x2 tables	1-sided p-value for larger tables
	n= 1215	%	n=62	%	n=59	%	n=6	%	OR	0.95	CI	p-value		
Age														
mean (sd)	38.2 (8.4)		38.8 (9.2)		39.4 (9.2)		44.0 (6.2)		1.01	0.98	1.04	0.618		
[range]	[18-67]		[20-62]		[21-62]		[39-56]							
HIV status														
Negative/Unknown	252	**20.74**	2	**3.23**	1	**1.69**	0	**0**	N/A				<0.001	
Positive	963	**79.26**	60	**96.77**	58	**98.31**	6	**100**						
Seen in a clinic in London														
No	394	**32.43**	12	**19.35**	10	**16.95**	0	**0**	1.00					
Yes	821	**67.57**	50	**80.65**	49	**83.05**	6	**100**	2.00	1.05	3.80	0.034		
Presentation day														
2010	847	**70.17**	54	**87.1**	25	**42.37**	1	**16.67**	1.00			0.006		
<2010	360	**29.83**	8	**12.9**	34	**57.63**	5	**83.33**	0.35	0.16	0.74			
Sexuality														
Homosexual	1,192	**98.11**	62	**100**	59	**100**	6	**100**	N/A				0.623	
Bisexual	23	**1.89**	0	**0**	0	**0**	0	**0**						
Ethnicity														
White	1,075	**88.48**	53	**85.48**	51	**86.44**	3	**50**	N/A					0.495
Black	53	**4.36**	3	**4.84**	3	**5.08**	1	**16.67**						
Asian	33	**2.72**	1	**1.61**	1	**1.69**	2	**33.33**						
Other	36	**2.96**	4	**6.45**	3	**5.08**	0	**0**						
Unknown	18	**1.48**	1	**1.61**	1	**1.69**	0	**0**						
Duration of symptoms														
Week or less	390	**32.1**	23	**37.1**	25	**42.37**	4	**66.67**	1.00					
More than a week	584	**48.07**	29	**46.77**	27	**45.76**	1	**16.67**	0.84	0.48	1.48	0.549		
Unknown	241.00	**19.84**	10.00	**16.13**	7.00	**11.86**	1.00	**16.67**	0.70	0.33	1.50	0.364		
Symptoms														
No	158	**13**	4	**6.45**	7	**11.86**	1	**16.67**	N/A					0.034
Yes	1,033	**85.02**	54	**87.1**	51	**86.44**	5	**83.33**						
Unknown	24.00	**1.98**	4	**6.45**	1	**1.69**	0	**0**						
Contact tracing														
No	1,097	**90.29**	53	**85.48**	53	**89.83**	6	**100**	1.00					0.071
Yes	94	**7.74**	5	**8.06**	5	**8.47**	0	**0**	1.10	0.43	2.82	0.841		
Unknown	24.00	**1.98**	4	**6.45**	1	**1.69**	0	**0**	N/A					
Routine STI screen														
No	1,104	**90.86**	55	**88.71**	56	**94.92**	5	**83.33**	N/A					0.079
Yes	87	**7.16**	3	**4.84**	2	**3.39**	1	**16.67**						
Unknown	24	**1.98**	4	**6.45**	1	**1.69**	0	**0**						
Referral														
No	1,144	**94.16**	58	**93.55**	58	**98.31**	6	**100**	N/A					0.023
Yes	47	**3.87**	0	**0**	0	**0**	0	**0**						
Unknown	24	**1.98**	4	**6.45**	1	**1.69**	0	**0**						
Location of symptoms reported														
None	63	**5.19**	3	**4.84**	4	**6.78**	1	**16.67**	N/A					0.995
Only Genital	60	**4.94**	2	**3.23**	3	**5.08**	0	**0**						
Only Rectal	817	**67.24**	44	**70.97**	39	**66.1**	5	**83.33**						
Both	204	**16.79**	10	**16.13**	12	**20.34**	0	**0**						
Unknown	71	**5.84**	3	**4.84**	1	**1.69**	0	**0**						
Site of infection														
Rectal	460	**37.86**	19	**30.65**	37	**62.71**	5	**83.33**	N/A					0.562
Genital	15	**1.23**	0	**0**	0	**0**	0	**0**						
Both or other (throat, n=1)	11	**0.91**	0	**0**	0	**0**	0	**0**						
Missing	729	**60**	43	**69.35**	22	**37.29**	1	**16.67**						

Any systemic symptom												
No	844	**69.47**	42	**67.74**	44	**74.58**	6	**100**	1.00			
Yes	319	**26.26**	18	**29.03**	12	**20.34**	0	**0**	1.13	0.64	2.00	0.664
Unknown	52	**4.28**	2	**3.23**	3	**5.08**	0	**0**	N/A			
No other STIs												
No	456	**37.53**	31	**50**	21	**35.59**	1	**16.67**	1.00			
Yes	691	**56.87**	27	**43.55**	35	**59.32**	5	**83.33**	0.57	0.34	0.98	0.040
Unknown	68	**5.6**	4	**6.45**	3	**5.08**	0	**0**	N/A			
Gonorrhoea												
No	950	**78.19**	40	**64.52**	44	**74.58**	5	**83.33**	1.00			
Yes	198	**16.3**	18	**29.03**	13	**22.03**	1	**16.67**	2.16	1.21	3.84	0.009
Unknown	67	**5.51**	4	**6.45**	2	**3.39**	0	**0**	N/A			
Syphilis												
No	1,078	**88.72**	54	**87.1**	52	**88.14**	6	**100**	1.00			
Yes	70	**5.76**	4	**6.45**	5	**8.47**	0	**0**	1.14	0.40	3.24	0.805
Unknown	67	**5.51**	4	**6.45**	2	**3.39**	0	**0**	N/A			
Any other STI^												
No	734	**60.41**	31	**50**	37	**62.71**	5	**83.33**	1.00			
Yes	346	**28.48**	23	**37.1**	17	**28.81**	1	**16.67**	1.57	0.90	2.73	0.109
Unknown	135	**11.11**	8	**12.9**	5	**8.47**	0	**0**	N/A			
Hepatitis C (PCR)												
No	422	**34.73**	21	**33.87**	23	**38.98**	1	**16.67**	1.00			
Yes	110	**9.05**	12	**19.35**	16	**27.12**	3	**50**	2.19	1.05	4.59	0.038
Unknown	683	**56.21**	29	**46.77**	20	**33.9**	2	**33.33**	0.85	0.48	1.52	0.588
Hepatitis C (Ab)												
No	800	**65.84**	35	**56.45**	29	**49.15**	1	**16.67**	1.00			
Yes	168	**13.83**	17	**27.42**	25	**42.37**	3	**50**	2.31	1.27	4.23	0.006
Unknown	247	**20.33**	10	**16.13**	5	**8.47**	2	**33.33**	0.93	0.45	1.90	0.832
Acquisition country												
UK	922	**75.88**	50	**80.65**	49	**83.05**	5	**83.33**	N/A			0.506
Abroad	92	**7.57**	2	**3.23**	2	**3.39**	1	**16.67**				
Either	55	**4.53**	4	**6.45**	1	**1.69**	0	**0**				
Unknown/ missing	146	**12.02**	6	**9.68**	7	**11.86**	0	**0**				
Met partners in any venue^												
None reported	306	**25.19**	17	**27.42**	15	**25.42**	2	**33.33**	1.00			
Met sex partners in these	274	**22.55**	18	**29.03**	19	**32.20**	2	**33.33**	1.18	0.60	2.34	0.630
Unknown/missing	635	**52.26**	27	**43.55**	25	**42.37**	2	**33.33**	0.77	0.41	1.43	0.399
Number of contacts												
median (range)	3 (0-213)		3 (0-85)		4 (0-100)		8.5 (3-20)		1.00	1.00	1.02	0.237
Receptive anal intercourse												
None reported	75	**6.17**	0	**0**	0	**0**	0	**0**	N/A			0.029
Reported protected/or unl	225	**18.52**	9	**14.52**	6	**10.17**	0	**0**	0.64	0.31	1.31	0.219
Unprotected	810	**66.67**	51	**82.26**	51	**86.44**	6	**100**	1.00			
Unknown	105	**8.64**	2	**3.23**	2	**3.39**	0	**0**	N/A			
Insertive anal intercourse												
None reported	109	**8.97**	4	**6.45**	9	**15.25**	3	**50**	N/A			0.380
Reported protected/or unl	204	**16.79**	9	**14.52**	6	**10.17**	0	**0**	0.71	0.34	1.48	0.361
Unprotected	642	**52.84**	40	**64.52**	36	**61.02**	3	**50**	1.00			
Unknown	260	**21.4**	9	**14.52**	8	**13.56**	0	**0**	0.56	0.27	1.16	0.118
Any fisting												
No fisting reported	485	**39.92**	26	**41.94**	21	**35.59**	1	**16.67**	N/A			0.639
Some fisting reported	54	**4.44**	3	**4.84**	4	**6.78**	0	**0**				
Both reported, unprotecte	64	**5.27**	5	**8.06**	7	**11.86**	1	**16.67**				
Some unknown	612	**50.37**	28	**45.16**	27	**45.76**	4	**66.67**				

Any oral sex													
None reported	98	**8.07**	1	**1.61**	1	**1.69**	0	**0**	N/A				0.301
Reported some	26	**2.14**	1	**1.61**	2	**3.39**	0	**0**					
Reported one unprotectec	38	**3.13**	1	**1.61**	1	**1.69**	0	**0**					
Reported both unrprotecte	869	**71.52**	51	**82.26**	51	**86.44**	6	**100**					
Some or all unknown	184	**15.14**	8	**12.9**	4	**6.78**	0	**0**					
Sharing sex toys													
No	478	**39.34**	23	**37.1**	17	**28.81**	2	**33.33**	1.00				
Any (prot or unpr)	79	**6.5**	9	**14.52**	6	**10.17**	0	**0**	2.37	1.06	5.30	0.036	
Unknown/ missing	658	**54.16**	30	**48.39**	36	**61.02**	4	**66.67**	0.95	0.54	1.65	0.849	
Vaginal intercourse													
No	903	**74.32**	52	**83.87**	38	**64.41**	6	**100**	N/A				0.323
Yes (unprot or not)	5	**0.41**	0	**0**	1	**1.69**	0	**0**					
Unknown/missing	307	**25.27**	10	**16.13**	20	**33.9**	0	**0**					

N/A, not applicable (parameter estimate not presented due to low number of events in the cell, and category excluded from the analysis.

Appendix for chapter 4

List of articles which were reviewed for the chapter. They are presented as they were originally divided at the early stages of the literature search (but might have subsequently been considered to belong to a different category).

#	Author,year	Title
"Structural"		
1	Bird2011a	Ethnic differences in HIV-disclosure and sexual risk
2	Bird2011	A conceptual model of HIV disclosure in casual sexual encounters among men who have sex with men
3	Carballo-Dieguez2006	Sexual negotiation, HIV-status disclosure, and sexual risk behavior among Latino men who use the Internet to seek sex with other men
4	Chaudoir2011	Understanding HIV disclosure: A review and application of the Disclosure Processes Model
5	Crepaz2009	Prevalence of unprotected anal intercourse among HIV-diagnosed MSM in the United States: a meta-analysis.
6	Daly2011	The role of perceived risk in general practitioners' decisions to inform partners of HIV-infected patients
7	DeWit2000	Understanding aids preventive behavior with casual and primary partners in homosexual men: The theory of planned behavior and the information-motivation-behavioral-skills model
8	Elford2008	Disclosure of HIV status - The role of ethnicity among people living with HIV in London
9	Espinosa2010	Risk-Taking Behaviors of Gay and Bisexual Men in New York City Post 9/11
10	Folch2005	Trends in the prevalence of HIV infection and risk behaviors in homo- and bisexual men
11	Hoff2012	Relationship Characteristics Associated with Sexual Risk Behavior Among MSM in Committed Relationships
12	Johnson2008a	Internalized Heterosexism Among HIV-Positive, Gay-Identified Men: Implications for HIV Prevention and Care
13	Johnson2008	Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men.
14	Kanter2011	Risk behaviour and HIV prevalence among men who have sex with men in a multiethnic society: a venue-based study in Kuala Lumpur, Malaysia
15	Lauby2008	Sexual risk behaviors of HIV-positive, HIV-negative, and serostatus-unknown Black men who have sex with men and women
16	Millett2012	Comparisons of disparities and risks of HIV infection in black and other men who have sex with men in Canada, UK, and USA: a meta-analysis
17	Munoz-Laboy2005	Sexual risk behaviour, viral load, and perceptions of HIV transmission among homosexually active Latino men: an exploratory study
18	Murphy2012	Seroadaptation in a Sample of Very Poor Los Angeles Area Men Who Have Sex with Men
19	O'Leary2007	Correlates of risk patterns and Race/Ethnicity among HIV-Positive men who have sex with men
20	Overstreet2012	Internalized stigma and HIV status disclosure among HIV-positive black men who have sex with men
21	Patel2012	Routine Brief Risk-Reduction Counseling With Biannual STD Testing Reduces STD Incidence Among HIV-Infected Men Who Have Sex With Men in Care
22	Poppen2004	Predictors of unprotected anal intercourse among HIV-positive Latino gay and bisexual men
23	Poppen2005	Serostatus disclosure, seroconcordance, partner relationship, and unprotected anal intercourse among HIV-positive Latino men who have sex with men
24	Rendina2012	Stigma and Sexual Compulsivity in a Community-Based Sample of HIV-Positive Gay and Bisexual Men
25	Stolte2006	Association between 'safer sex fatigue' and rectal gonorrhea is mediated by unsafe sex with casual partners among HIV-positive homosexual men
26	Tieu2011	Sexual partner characteristics, serodiscordant/serostatus unknown unprotected anal intercourse and disclosure among human immunodeficiency virus-infected and uninfected black men who have sex with men in New York City
27	Truong2006	Increases in sexually transmitted infections and sexual risk behaviour without a concurrent increase in HIV incidence among men who have sex with men in San Francisco: a suggestion of HIV serosorting?
28	Wei2012	HIV Disclosure and Sexual Transmission Behaviors among an Internet Sample of HIV-positive Men Who Have Sex with Men in Asia: Implications for Prevention with Positives
29	Wei2011a	Racial/Ethnic differences in seroadaptive and serodisclosure behaviors among men who have sex with men.
30	Woodward2010	Computer-assisted survey of attitudes to HIV and sexually transmissible infection partner notification in HIV-positive men who have sex with men
31	Xia2006	The effect of venue sampling on estimates of HIV prevalence and sexual risk behaviors in men who have sex with men
32	Zea2003	Asking and telling: Communication about HIV status among Latino HIV-positive gay men
"Community"		
33	Adam2008	Silence, assent and HIV risk
34	Benotsch2006	Substance use, medications for sexual facilitation, and sexual risk behavior among traveling men who have sex with men
35	Berry2008	The Internet, HIV serosorting and transmission risk among men who have sex with men, San Francisco.
36	Bianchi2006	A comparison of HIV-related factors among seropositive Brazilian, South American, and Puerto Rican gay men in the United States
37	Bolding2005	Gay men who look for sex on the Internet: is there more HIV/STI risk with online partners?
38	Bolding2004	Use of gay Internet sites and views about online health promotion among men who have sex with men
39	Brewer2006	Unsafe sexual behavior and correlates of risk in a probability sample of men who have sex with men in the era of highly active antiretroviral therapy
40	Chen2012a	Prevalence and Predictors of Conscious Risk Behavior Among San Franciscan Men who have Sex with Men
41	Clatts2005	An emerging HIV risk environment: a preliminary epidemiological profile of an MSM POZ Party in New York City.
42	Courtenay-Quirk2008	Factors associated with sexual risk behavior among persons living with HIV: Gender and sexual identity group differences
43	Courtenay-Quirk2003	Interests in HIV prevention topics of HIV-seropositive men who have sex with men
44	Courtenay-Quirk2006	Is HIV/AIDS stigma dividing the gay community? Perceptions of HIV-positive men who have sex with men
45	Cox2004	HIV status of sexual partners is more important than antiretroviral treatment related perceptions for risk taking by HIV positive MSM

in Montreal, Canada.

46 Darbes2012 Partner-provided social support influences choice of risk reduction strategies in gay male couples.
47 Edwards-Jackson2012 HIV serostatus disclosure is not associated with safer sexual behavior among HIV-positive men who have sex with men (MSM) and their partners at risk for infection in Bangkok, Thailand
48 Elford2004 Web-based behavioral surveillance among men who have sex with men: a comparison of online and offline samples in London, UK
49 Elford2001 Seeking sex on the Internet and sexual risk behaviour among gay men using London gyms
50 Fan2012 Behavioral and Serologic Survey of Men Who Have Sex with Men in Beijing, China: Implication for HIV Intervention
51 Fisher2006 Recruitment source and behavioral risk patterns of HIV-positive men who have sex with men
52 Forrest2010 Crystal Methamphetamine Use and Sexual Risk Behaviors among HIV-Positive and HIV-Negative Men Who Have Sex with Men in South Florida
53 Golden2007 Ongoing risk behavior among persons with HIV in medical care.
54 Gorbach2009 Bridging Sexual Boundaries: Men Who Have Sex with Men and Women in a Street-Based Sample in Los Angeles
55 Grov2007 Barebacking, the Internet, and harm reduction: an intercept survey with gay and bisexual men in Los Angeles and New York City.
56 Grov2010a HIV STATUS DIFFERENCES IN VENUES WHERE HIGHLY SEXUALLY ACTIVE GAY AND BISEXUAL MEN MEET SEX PARTNERS: RESULTS FROM A PILOT STUDY
57 Hatala1998 Dating with HIV: A content analysis of gay male HIV-positive and HIV-negative personal advertisements
58 Hospers2005 A new meeting place: chatting on the Internet, e-dating and sexual risk behaviour among Dutch men who have sex with men
59 Leobon2011 A relative profile of HIV-negative users of French websites for men seeking men and predictors of their regular risk taking: a comparison with HIV-positive users
60 Liau2006 Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men.
61 Mimiaga2010 It's a Quick Way to Get What You Want": A Formative Exploration of HIV Risk Among Urban Massachusetts Men Who Have Sex with Men Who Attend Sex Parties"
62 Mimiaga2011 Sex Parties among Urban MSM: An Emerging Culture and HIV Risk Environment
63 Mimiaga2009 Street Workers and Internet Escorts: Contextual and Psychosocial Factors Surrounding HIV Risk Behavior among Men Who Engage in Sex Work with Other Men
64 Miner2009 How Do Social Norms Impact HIV Sexual Risk Behavior in HIV-Positive Men Who Have Sex with Men? Multiple Mediator Effects
65 Pollock2009 Environmental factors in relation to unprotected sexual behavior among gay, bisexual, and other MSM
66 VanderBij2007 Condom use rather than serosorting explains differences in HIV incidence among men who have sex with men.
67 Velter2009 Do men who have sex with men use serosorting with casual partners in France? Results of a nationwide survey (ANRS-EN17-Presse Gay 2004).
68 Wei2012a Sexual Transmission Behaviors and Serodiscordant Partnerships Among HIV-Positive Men Who Have Sex With Men in Asia
69 Williamson2006 Increases in HIV-related sexual risk behavior among community samples of gay men in London and Glasgow: how do they compare?
70 Winter2012 Discussion of HIV Status by Serostatus and Partnership Sexual Risk Among Internet-Using MSM in the United States
71 Zablotska2008 Differing trends in sexual risk behaviours in three Australian states: New South Wales, Victoria and Queensland, 1998-2006

"Individuals and partnerships"

72 Bachmann2009 Risk behaviours in HIV-positive men who have sex with men participating in an intervention in a primary care setting
73 Balthasar2010 Intentional Risk Reduction Practices of Men in Switzerland Who Have Anal Intercourse With Casual Male Partners
74 Bouhnik2007 Unprotected sex in regular partnerships among homosexual men living with HIV: a comparison between sero-nonconcordant and seroconcordant couples (ANRS-EN12-VESPA Study)
75 Bourne2011 Perceptions of superinfection risk among gay men with diagnosed HIV who have unprotected anal intercourse
76 Bousman2009 Negative mood and sexual behavior among non-monogamous men who have sex with men in the context of methamphetamine and HIV
77 Brooks2011 Motivators, concerns, and barriers to adoption of preexposure prophylaxis for HIV prevention among gay and bisexual men in HIV-serodiscordant male relationships
78 Bruce2012 Sexual Risk Behavior and Risk Reduction Beliefs Among HIV-Positive Young Men Who have Sex with Men.
79 Burt2012 Variations in Patterns of Sexual Risk Behavior among Seattle-Area MSM Based on their HIV Status, the HIV Status of their Partner and Partner Type
80 Campbell2009 HIV-1 Superinfection in the Antiretroviral Therapy Era: Are Seroconcordant Sexual Partners at Risk?
81 Chen2003 Unprotected anal intercourse between potentially HIV-serodiscordant men who have sex with men, San Francisco.
82 Davis2011 Love and HIV serodiscordance in gay men's accounts of life with their regular partners
83 Denning2005 Unprotected anal intercourse among HIV-positive men who have a steady male sex partner with negative or unknown HIV serostatus
84 DeRosa1998 Preventive counseling of HIV-positive men and self-disclosure of serostatus to sex partners: New opportunities for prevention
85 Dodds2004 Increasing risk behaviour and high levels of undiagnosed HIV infection in a community sample of homosexual men
86 Dolcini2003 The HIV epidemic among older men who have sex with men
87 Dolezal2005 Alcohol, marijuana, cocaine use, and relationship quality among HIV serodiscordant male couples
88 Dombrowski2013 Evidence for the Long-Term Stability of HIV Transmission-Associated Sexual Behavior After HIV Diagnosis.
89 Dombrowski2010 Antiretroviral Therapy Use Among HIV-Infected Men Who Have Sex With Men Attending a Sexually Transmitted Diseases Clinic
90 Dow1996 An empirical examination of the implications of assortative matching on the incidence of HIV
91 Dubois-Arber2012 Risk reduction practices in men who have sex with men in Switzerland: serosorting, strategic positioning, and withdrawal before ejaculation.
92 Eaton2009 HIV transmission risk among HIV seroconcordant and serodiscordant couples: dyadic processes of partner selection.
93 Elford2007 Barebacking among HIV-positive gay men in London
94 Elford2004a Trends in sexual behaviour among London homosexual men 1998-2003: implications for HIV prevention and sexual health promotion
95 Elford2000 Combination therapies for HIV and sexual risk behavior among gay men
96 Elford2002 High-risk sexual behaviour increases among London gay men between 1998 and 2001: what is the role of HIV optimism?
97 Elford2007a Sexual behaviour of people living with HIV in London: implications for HIV transmission
98 Eustace2010 HIV disclosure among HIV positive individuals: a concept analysis
99 Fernandez-Davila2011 Silence and Assumptions: Narratives on the Disclosure of HIV Status to Casual Sexual Partners and Serosorting in a Group of Gay Men in Barcelona
100 Frost2008 Understanding why gay men seek HIV-seroconcordant partners: intimacy and risk reduction motivations.
101 Golden2004 Importance of sex partner HIV status in HIV risk assessment among men who have sex with men.
102 Gorbach2004 Don't ask, don't tell: patterns of HIV disclosure among HIV positive men who have sex with men with recent STI practising high risk behaviour in Los Angeles and Seattle
103 Gorbach2006 Transmission behaviors of recently HIV-infected men who have sex with men
104 Gorbach2011 Behaviors of recently HIV-infected men who have sex with men in the year postdiagnosis: effects of drug use and partner types
105 Guzman2006 Communication of HIV viral load to guide sexual risk decisions with serodiscordant partners among San Francisco men who have sex with men
106 Halkitis2005 Seroconcordant sexual partnerings of HIV-seropositive men who have sex with men
107 Halkitis2003 Barebacking among gay and bisexual men in New York City: Explanations for the emergence of intentional unsafe behavior

108	Harawa2006	Perceptions towards condom use, sexual activity, and HIV disclosure among HIV-positive African American men who have sex with men: Implications for heterosexual transmission
109	Hart2005	Partner awareness of the serostatus of HIV-seropositive men who have sex with men: impact on unprotected sexual behavior.
110	Hatfield2009	Comparison of Substance Use and Risky Sexual Behavior Among a Diverse Sample of Urban, HIV-Positive Men Who Have Sex with Men
111	Hoff2009	SEROSTATUS DIFFERENCES AND AGREEMENTS ABOUT SEX WITH OUTSIDE PARTNERS AMONG GAY MALE COUPLES
112	Hoff1997	Differences in sexual behavior among HIV discordant and concordant gay men in primary relationships
113	Holt2013	The converging and diverging characteristics of HIV-positive and HIV-negative gay men in the Australian Gay Community Periodic Surveys, 2000-2009
114	Holt2011	Predictors of HIV Disclosure Among Untested, HIV-Negative and HIV-Positive Australian Men Who had Anal Intercourse with Their Most Recent Casual Male Sex Partner
115	Holt1998	The role of disclosure in coping with HIV infection
116	Horvath2008	Sexual negotiation and HIV serodisclosure among men who have sex with men with their online and offline partners
117	Horvath2012	Discussions of viral load in negotiating sexual episodes with primary and casual partners among men who have sex with men
118	Houston2012	Depressive Symptoms Among MSM Who Engage in Bareback Sex: Does Mood Matter?
119	Klitzman2007	It's not just what you say: Relationships of HIV dislosure and risk reduction among MSM in the post-HAART era
120	Kurtz2012	Resilience, syndemic factors, and serosorting behaviors among HIV-positive and HIV-negative substance-using MSM.
121	Lampe2012	ART use, viral suppression, and sexual behaviour among HIV-diagnosed MSM in the UK: results from the Antiretrovirals, Sexual Transmission Risk and Attitudes (ASTRA) Study
122	Larkins2005	Methamphetamine-dependent gay men's disclosure of their HIV status to sexual partners
123	Lattimore2011	Changing patterns of sexual risk behavior among London gay men: 1998-2008.
124	Lightfoot2005	The influence of partner type and risk status on the sexual behavior of young men who have sex with men living with HIV/AIDS
125	Mao2011	Rates of condom and non-condom-based anal intercourse practices among homosexually active men in Australia: deliberate HIV risk reduction?
126	Marcus2011	HIV serosorting among HIV-positive men who have sex with men is associated with increased self-reported incidence of bacterial sexually transmissible infections.
127	Marks2001	HIV-positive men's sexual practices in the context of self-disclosure of HIV status
128	Marks1991	Self-disclosure of HIV infection to sexual partners.
129	Marks1994	ANAL INTERCOURSE AND DISCLOSURE OF HIV-INFECTION AMONG SEROPOSITIVE GAY AND BISEXUAL MEN
130	McConnell2010	Sexual seroadaptation: lessons for prevention and sex research from a cohort of HIV-positive men who have sex with men
131	McCready2008	HIV serostatus disclosure to sexual partners among HIV-positive methamphetamine-using gay, bisexual, and other men who have sex with men
132	McDaid2012	Serosorting and strategic positioning during unprotected anal intercourse: are risk reduction strategies being employed by gay and bisexual men in Scotland?
133	McFarland2012	Behavior, intention or chance? A longitudinal study of HIV seroadaptive behaviors, abstinence and condom use.
134	McFarland2011	HIV seroadaptation among individuals, within sexual dyads, and by sexual episodes, men who have sex with men, San Francisco, 2008
135	Menza2011	Stable Sexual Risk Behavior in a Rapidly Changing Risk Environment: Findings from Population-Based Surveys of Men Who Have Sex with Men in Seattle, Washington, 2003-2006
136	Moreau-Gruet2001	Management of the risk of HIV infection in male homosexual couples
137	Morin2008	A behavioral intervention reduces HIV transmission risk by promoting sustained serosorting practices among HIV-infected men who have sex with men.
138	Moskowitz2008	Vengeance, HIV disclosure, and perceived HIV transmission to others
139	Moskowitz2011	Self-Esteem in HIV-Positive and HIV-Negative Gay and Bisexual Men: Implications for Risk-Taking Behaviors with Casual Sex Partners
140	O'Dell2008	HIV prevention altruism and sexual risk behavior in HIV-Positive men who have sex with men
141	Ostrow2008	Prospective study of attitudinal and relationship predictors of sexual risk in the multicenter AIDS cohort study
142	Owen2008	An 'elephant in the room'? Stigma and hepatitis C transmission among HIV-positive 'serosorting' gay men.
143	Pantalone2008	Motivations for the recreational use of erectile enhancing medications in urban gay and bisexual men
144	Parsons2007	Intentional unprotected anal intercourse among sex who have sex with men: Barebacking-from behavior to identity
145	Parsons2005a	Consistent, inconsistent, and non-disclosure to casual sexual partners among HIV-seropositive gay and bisexual men.
146	Parsons2005	Sexual harm reduction practices of HIV-seropositive gay and bisexual men: serosorting, strategic positioning, and withdrawal before ejaculation.
147	Parsons2006	Positive, negative, unknown: assumptions of HIV status among HIV-positive men who have sex with men.
148	Patel2006	Circuit parties: sexual behaviors and HIV disclosure practices among men who have sex with men at the White Party, Palm Springs, California, 2003.
149	Prestage2001	Gay men's casual sex encounters: discussing HIV and using condoms
150	Przybyla2012	Serostatus disclosure to sexual partners among people living with HIV: Examining the roles of partner characteristics and stigma
151	Purcell2005	Illicit substance use, sexual risk, and HIV-positive gay and bisexual men: differences by serostatus of casual partners.
152	Rajasingham2012	A Systematic Review of Behavioral and Treatment Outcome Studies Among HIV-Infected Men Who Have Sex with Men Who Abuse Crystal Methamphetamine
153	Rawstorne2007a	Differences between HIV-positive gay men who 'frequently', 'sometimes' or 'never' engage in unprotected anal intercourse with serononconcordant casual partners: Positive Health cohort, Australia
154	Relf2009	A QUALITATIVE ANALYSIS OF PARTNER SELECTION, HIV SEROSTATUS DISCLOSURE, AND SEXUAL BEHAVIORS AMONG HIV-POSITIVE URBAN MEN
155	Remien2003	Levels and correlates of psychological distress in male couples of mixed HIV status
156	Ross2008	THE RELATIONSHIP OF INTERNALIZED HOMONEGATIVITY TO UNSAFE SEXUAL BEHAVIOR IN HIV-SEROPOSITIVE MEN WHO HAVE SEX WITH MEN
157	Rosser2010	Effects of a behavioral intervention to reduce serodiscordant unsafe sex among HIV positive men who have sex with men: the Positive Connections randomized controlled trial study
158	Rosser2008	Predictors of HIV disclosure to secondary partners and sexual risk behavior among a high-risk sample of HIV-positive MSM: results from six epicenters in the US
159	Rutledge2009	Formation of Personal HIV Disclosure Policies among HIV-Positive Men Who Have Sex with Men
160	Schutz2011	Determinants of condom use among HIV-positive men who have sex with men
161	Schwarcz2007	Prevalence of HIV infection and predictors of high-transmission sexual risk behaviors among men who have sex with men
162	Semple2006	Sexual risk behavior of HIV-positive methamphetamine-using men who have sex with men: The role of partner serostatus and partner type
163	Serovich2005	Methods of HIV disclosure by men who have sex with men to casual sexual partners
164	Sheon2004	Ambivalent tales of HIV disclosure in San Francisco
165	Sirivongrangson2012	HIV Risk Behavior Among HIV-Infected Men Who have Sex with Men in Bangkok, Thailand
166	Snowden2011	Seroadaptive behaviours among men who have sex with men in San Francisco: the situation in 2008.

167	Snowden2009	Prevalence of seroadaptive behaviours of men who have sex with men, San Francisco, 2004.
168	Spindler2007	Viagra, methamphetamine, and HIV risk: results from a probability sample of MSM, San Francisco
169	Stall2011	Do rates of unprotected anal intercourse among HIV-positive MSM present a risk for hepatitis C transmission?
170	Stolte2004	Perceived viral load, but not actual HIV-1-RNA load, is associated with sexual risk behaviour among HIV infected homosexual men
171	Suzan-Monti2011	The burden of HIV experience and care among MSM having an HIV-positive seroconcordant steady partner: a possible research hypothesis. Results from the French VESPA ANRS EN-12 study
172	Theodore2004	Intimacy and sexual behavior among HIV-positive men-who-have-sex-with-men in primary relationships
173	Vanable2012	What HIV-Positive MSM Want from Sexual Risk Reduction Interventions: Findings from a Qualitative Study
174	VandeVen2005	Undetectable viral load is associated with sexual risk taking in HIV serodiscordant gay couples in Sydney
175	VandeVen2002a	In a minority of gay men, sexual risk practice indicates strategic positioning for perceived risk reduction rather than unbridled sex.
176	Wiktor1990	EFFECT OF KNOWLEDGE OF HUMAN IMMUNODEFICIENCY VIRUS-INFECTION STATUS ON SEXUAL-ACTIVITY AMONG HOMOSEXUAL MEN
177	Williamson2008	Sexual risk behaviour and knowledge of HIV status among community samples of gay men in the UK.
178	Wolitski2005	Prevention with gay and bisexual men living with HIV: rationale and methods of the Seropositive Urban Men's Intervention Trial (SUMIT)
179	Wolitski2004	Prevention with HIV-seropositive men who have sex with men - Lessons from the Seropositive Urban Men's Study (SUMS) and the Seropositive Urban Men's Intervention Trial (SUMIT)
180	Wolitski1998	HIV serostatus disclosure among gay and bisexual men in four American cities: general patterns and relation to sexual practices
181	Wolitski2003	Self-perceived responsibility of HIV-Seropositive men who have sex with men for preventing HIV transmission
182	Wolitski2005a	Effects of a peer-led behavioral intervention to reduce HIV transmission and promote serostatus disclosure among HIV-seropositive gay and bisexual men
183	Xia2006a	Knowledge of sexual partner's HIV serostatus and serosorting practices in a California population-based sample of men who have sex with men.
184	Zablotska2009	Gay men's current practice of HIV seroconcordant unprotected anal intercourse: serosorting or seroguessing?
185	Zekan2008	Unsafe sexual behavior among HIV-infected patients in Croatia, 2006: Prevalence and associated factors

Found through reference lists and Sigma research

187	Vanable(2003)	Viral load and HIV treatment attitudes as correlates of sexual risk behavior among HIV-positive gay men
188	Parsons (2003)	Correlates of sexual risk behaviors among HIV-positive men who have sex with men.
189	Dodds (2009)	Responses to criminal prosecutions for HIV transmission among gay men with HIV in England and Wales.
190	Chlamers (2002)	The criminalisation of HIV transmission
191	Vallabhaneni	Changes in Seroadaptive Practices from before to after Diagnosis of Recent HIV Infection among Men Who Have Sex with Men.
192	Byrne (2013)	Nondisclosure prosecutions and population health outcomes: examining HIV testing, HIV diagnoses, and the attitudes of men who have sex with men following nondisclosure prosecution media releases in Ottawa, Canada
193	Brennan (2010)	HIV treat optimish and UAI in HIVpos MSM
194	Flowers (2000)	Community, responsibility and culpability: HIV risk-management among Scottish gay men.
195	Davies (Sigma)	Putting it about: Health promotion for gay men with higher numbers of sexual partners
196	Dodds (Sigma)	A telling dilemma HIV disclosure between male (homo)sexual partners
197	Dodds (Sigma)	Sexually charged: The views of gay and bisexual men on criminal prosecutions for sexual HIV transmission
198	Keogh (Sigma)	Relative safety Risk and unprotected anal intercourse among gay men diagnosed with HIV
199	Bourne (Sigma)	Relative safety II: Risk and unprotected anal intercourse among gay men with diagnosed HIV

Appendix for Chapter 5

Estimating the partner change rate and proportion of population in the highest activity group

Table S 1. Comparison of partner change rates in Sigma research, Natsal 2000 and LGV Enhanced Surveillance, categorised as in Sigma research

Sigma London (2008) UAI partners (12months)		Sigma London (2008) partners[1] (12 months)		MSM Natsal 2000 partners[2] (12 months)		LGV ES partners (2004-2010) (3 months)		GRASP Surveillance (2010); partners[1] (3 months)	
n=2372		n=2372		n=148		n=1189		n=568	
Partners	%	Partners	%	Partners	%	Partners	%	Partners	%
0	47.7%	0	4.5%	0	21.6%	0	1.9%		
1	28.0%	1	16.0%	1	32.4%	1	19.8%	0-1	24.8%
2	9.2%	2-4	24.4%	2	7.4%	2	23.6%		
3-4	7.1%			3-4	17.6%	3-4	18.3%	2-5	51.8%
+5	7.9%	5+	55.0%	5+	21.0%	+5	30.4%	+5	23.4%

Table S 2. Partnerships in the datasets which are above five. Data is presented as reported and for LGV Enhanced Surveillance the data is grouped as in Sigma London. Percentage values are proportions of the whole sample.

Sigma London (2008) UAI partners (12months)		Sigma London (2008) partners (12 months)		MSM Natsal 2000 partners (12 months)		LGV Enhanced surveillance partners (2004-2010) (3 months)		GRASP Surveillance (2010); partners (3 months)	
N=2372^		N=2372^		N=148		N=1189		n=568	
Partners	%	Partners	%	Partners	%	Partners	%	Partners	%
+5	7.9 %	5+	55.0%	5+	21.0%	+5	30.4%	+5	23.4%
		5-12	24.9%	5-10	13.5%	5-12	22.8%	6-10	14.8%
		13-29	14.7%	15-25	4.1%	13-29	7.6%	11+	8.6%
		30-	15.4%	55-100	3.4%	30-213	6.0%		

1. Sigma 2008, London summary report (Sigma Research, 2009)

   Number of men who responded to the survey in London; no information available how many answered the question on partner numbers.

2. Natsal 2000 (information obtained from Dr Peter White)

   Base: Men aged 16-44 who reported ≥1 male partners in the past 5 years in Natsal 2000

3. Gonorrhoea resistance antimicrobial Surveillance programme (GRASP), information for MSM partners reported in 2010; summary numbers received from Public Health England (Dr Gwenda Hughes, personal communication).

In Sigma (2008) London report the partner number for the past 12 months is higher than for Natsal with 55% reporting 5 or more partners compared to 21% (Table S1). To extrapolate from the data available we look at the distribution of those with 5 or more partners to get an estimate of the proportion potentially belonging to high-activity group and what their partner change rates may be (Table S2).

National Enhanced Syphilis Surveillance (NESS) for 1999 to 2008 with 8656 MSM. In total they reported a median 2 (interquartile range 2 to 5) sex partners in the past 3 months (Jebbari, Simms, Conti, *et al.*, 2011)

Table S 3. Description of how the Sigma 2008 London summary data on number of UAI partners in the past 12 months was used to estimate the range of partner change rate in the model as well as proportion of individuals belonging to low, medium and high risk categories.

SIGMA data (London 2008)		**"Calculations"**			**Model estimates**						**Population size**
Range	Prop	Prop/weight within group	pcr mid-point	weighted median	Acitvity class	pcr range	Proportion in group	pcr /year	pcr /month	pcr range /month	70000
None	47.70%	**63.01%**	0	0							
One	28.00%	**36.99%**	1	0.37	**Low risk**	0-1	76%	0.37	**0.03**	0.03-0.25	**52990**
Two	9.20%	**43.40%**	2	0.87							
3-4	7.10%	**33.49%**	3.5	1.17	**Medium risk**	2-5	21%	3.2	**0.27**	0.26-1.06	**14840**
5	4.90%	**23.11%**	5	1.16							
5+	3.00%	**100.00%**	"		**High-risk**	10-??	3%	13	**1.08**	1.07-1.7	**2100**

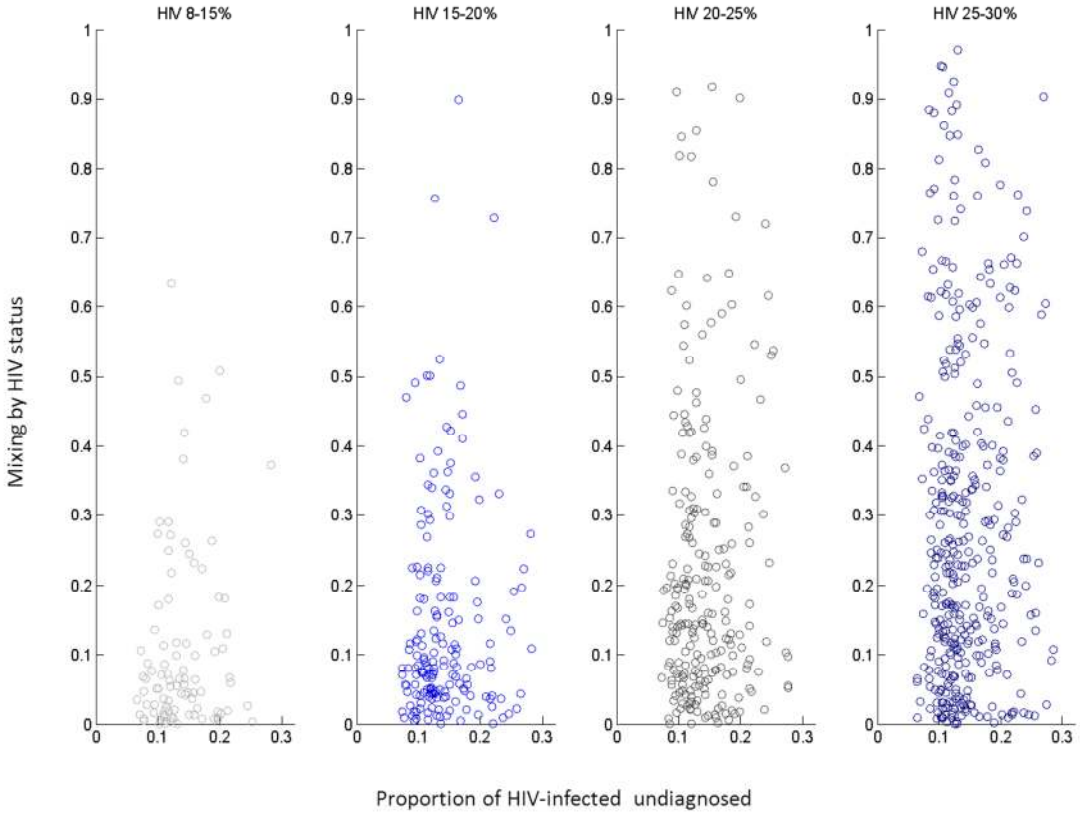


Figure S 1. Scatter plots of parameter controlling for serosorting mixing and proportion of HIV-infected who are undiagnosed when HIV-prevalence is in equilibrium.

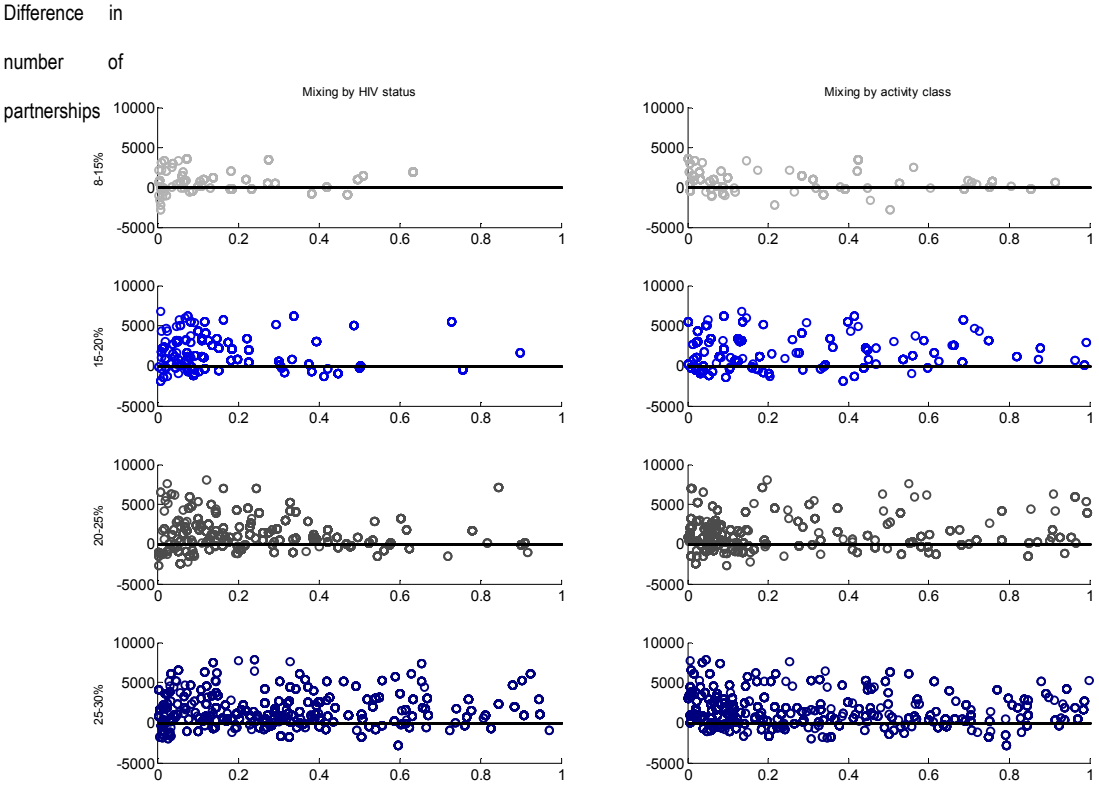


Figure S 2. Dispersion of partnerships in 1-5% LGV prevalence settings across

References

Adam, B.D., Husbands, W., Murray, J. & Maxwell, J. (2008) Silence, assent and HIV risk. *Culture Health & Sexuality*. 10 (8), 759–772.

Aghaizu, A., Mercey, D., Copas, A., Johnson, A.M., *et al.* (2013) Who would use PrEP? Factors associated with intention to use among MSM in London: a community survey. *Sexually transmitted infections*. 89 (3), 207–211.

Agresti, A. (2002) *Categorical Data Analysis*. 2nd Ed.. Hoboken, NJ, USA, John Wiley & Sons, Inc.

Ahdoot, A., Kotler, D.P., Suh, J.S., Kutler, C., *et al.* (2006) Lymphogranuloma venereum in human immunodeficiency virus-infected individuals in New York City. *Journal of clinical gastroenterology*. 40 (5), 385–390.

Alexander, S. (2009) The challenges of detecting gonorrhoea and chlamydia in rectal and pharyngeal sites: could we, should we, be doing more? *Sexually Transmitted Infections*. 85 (3), 159–160.

Al-Mutairi, N., Joshi, A., Nour-Eldin, O., Sharma, A.K., *et al.* (2007) Clinical patterns of sexually transmitted diseases, associated sociodemographic characteristics, and sexual practices in the Farwaniya region of Kuwait. *International journal of dermatology*. 46 (6), 594–599.

Altman, D.G., Machin, D., Bryant, T.N. & Gardner, J.D. (2000) *Statistics with confidence: Confidence intervals and statistical guidelines*. 2nd Ed. BMJ Books.

Anderson, A. & May, R. (1991) *Infectious Diseases of Humans: Dynamics and Control*. New York, Oxford University Press.

Annamunthodo, H. (1961) Rectal lymphogranuloma venereum in Jamaica. *Diseases of the Colon & Rectum*. 4 (1), 17–26.

Annan, N.T., Sullivan, A.K., Nori, A., Naydenova, P., *et al.* (2009) Rectal chlamydia--a reservoir of undiagnosed infection in men who have sex with men. *Sexually Transmitted Infections*. 85 (3), 176–179.

Aral, S.O. (2002) Determinants of STD epidemics: implications for phase appropriate intervention strategies. *Sexually Transmitted Infections*. 78 Suppl 1i3–13.

Bachmann, L.H., Grimley, D.M., Chen, H., Aban, I., *et al.* (2009) Risk behaviours in HIV-positive men who have sex with men participating in an intervention in a primary care setting. *International Journal of STD & AIDS*. 20 (9), 607–612.

Baggaley, R.F., White, R.G. & Boily, M.-C. (2010) HIV transmission risk through anal intercourse: systematic review, meta-analysis and implications for HIV prevention. *International journal of epidemiology*. 39 (4), 1048–1063.

Batteiger, B.E., Tu, W., Ofner, S., Van Der Pol, B., *et al.* (2010) Repeated Chlamydia trachomatis genital infections in adolescent women. *The Journal of infectious diseases*. 201 (1), 42–51.

Bauwens, J.E., Orlander, H., Gomez, M.P., Lampe, M., *et al.* (2002) Epidemic Lymphogranuloma venereum during epidemics of crack cocaine use and HIV infection in the Bahamas. *Sexually Transmitted Diseases*. 29 (5), 253–259.

Berg, R.C. (2009) Barebacking: a review of the literature. *Archives of sexual behavior*. 38 (5), 754–764.

Berglund, T., Bratt, G., Herrmann, B., Karlsson, A., *et al.* (2005) Two cases of lymphogranuloma venereum (LGV) in homosexual men in Stockholm. *Euro Surveillance*. 10 (3).

Berry, M., Raymond, H.F., Kellogg, T. & McFarland, W. (2008) The Internet, HIV serosorting and transmission risk among men who have sex with men, San Francisco. *AIDS (London, England)*. 22 (6), 787–789.

Beyrer, C., Baral, S.D., Van Griensven, F., Goodreau, S.M., *et al.* (2012) Global epidemiology of HIV infection in men who have sex with men. *Lancet*. 380 (9839), 367–377.

Van der Bij, A.K., Spaargaren, J., Morre, S.A., Fennema, H.S., *et al.* (2006) Diagnostic and clinical implications of anorectal lymphogranuloma venereum in men who have sex with men: a retrospective case-control study. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 42 (2), 186–194.

Bird, J.D.P., Fingerhut, D.D. & McKirnan, D.J. (2011) Ethnic differences in HIV-disclosure and sexual risk. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv*. 23 (4), 444–448.

Bird, J.D.P. & Voisin, D.R. (2011) A conceptual model of HIV disclosure in casual sexual encounters among men who have sex with men. *Journal of Health Psychology*. 16 (2), 365–373.

Birrell, P.J., Gill, O.N., Delpech, V.C., Brown, A.E., *et al.* (2013) HIV incidence in men who have sex with men in England and Wales 2001–10: a nationwide population study. *The Lancet Infectious Diseases*.

Blanchard, J.F. (2002) Populations, pathogens, and epidemic phases: closing the gap between theory and practice in the prevention of sexually transmitted diseases. *Sexually Transmitted Infections*. 78 Suppl 1i183–8.

Blower, S.M. & Dowlatabadi, H. (1994) Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example. *International Statistical Review / Revue Internationale de Statistique*. 62 (2), 229–243.

Boily, M.-C., Godin, G., Hogben, M., Sherr, L., *et al.* (2005) The impact of the transmission dynamics of the HIV/AIDS epidemic on sexual behaviour: a new hypothesis to explain recent increases in risk taking-behaviour among men who have sex with men. *Medical hypotheses*. 65 (2), 215–226.

Bolding, G., Davis, M., Hart, G., Sherr, L., *et al.* (2005) Gay men who look for sex on the Internet: is there more HIV/STI risk with online partners? *AIDS (London, England)*. 19 (9), 961–968.

Borg, M.L., Modi, A., Tostmann, A., Gobin, M., *et al.* (2012) Ongoing outbreak of Shigella flexneri serotype 3a in men who have sex with men in England and Wales, data from 2009-2011. *Euro surveillance*. 17 (13).

Bourne, A., Dodds, C., Keogh, P., Weatherburn, P., *et al.* (2009) *Relative safety II*. p.36.

Bradshaw, D., Matthews, G. & Danta, M. (2013) Sexually transmitted hepatitis C infection: the new epidemic in MSM? *Current opinion in infectious diseases*. 26 (1), 66–72.

Broadbent, A., Horner, P., Wills, G., Ling, A., *et al.* (2011) HIV-1 does not significantly influence Chlamydia trachomatis serovar L2 replication in vitro. *Microbes and infection / Institut Pasteur*. 13 (6), 575–584.

Brook, G., McSorley, J. & Shaw, A. (2013) Retrospective study of the effect of enhanced systematic sexually transmitted infection screening, facilitated by the use of electronic patient records, in an HIV-infected cohort. *HIV medicine*.

Brunham, R.C. (2013) Immunity to Chlamydia trachomatis. *The Journal of infectious diseases*. [Online] Available from: doi:10.1093/infdis/jit095.

Cassels, S. & Goodreau, S.M. (2011) Interaction of mathematical modeling and social and behavioral HIV/AIDS research. *Current opinion in HIV and AIDS*. 6 (2), 119–123.

Chalmers, J. (2002) The criminalisation of HIV transmission. *Journal of Medical Ethics*. 28 (3), 160–163.

Chaudoir, S.R., Fisher, J.D. & Simoni, J.M. (2011) Understanding HIV disclosure: A review and application of the Disclosure Processes Model. *Social Science & Medicine*. 72 (10), 1618–1629.

Chen, C.Y., Chi, K.H., Alexander, S., Ison, C.A., *et al.* (2008) A real-time quadriplex PCR assay for the diagnosis of rectal lymphogranuloma venereum and non-lymphogranuloma venereum Chlamydia trachomatis infections. *Sexually Transmitted Infections*. 84 (4), 273–276.

Chen, C.Y., Chi, K.H., Alexander, S., Martin, I.M., *et al.* (2007) The molecular diagnosis of lymphogranuloma venereum: evaluation of a real-time multiplex polymerase chain reaction test using rectal and urethral specimens. *Sexually Transmitted Diseases*. 34 (7), 451–455.

Chen, Y.H., Raymond, H.F., Grasso, M., Nguyen, B., *et al.* (2012) Prevalence and Predictors of Conscious Risk Behavior Among San Franciscan Men who have Sex with Men. *AIDS and behavior.*

Choudhury, B., Risley, C.L., Ghani, A.C., Bishop, C.J., *et al.* (2006) Identification of individuals with gonorrhoea within sexual networks: a population-based study. *Lancet*. 368 (9530), 139–146.

Christerson, L., Bom, R.J.M., Bruisten, S.M., Yass, R., *et al.* (2012) Chlamydia trachomatis Strains Show Specific Clustering for Men Who Have Sex with Men Compared to Heterosexual Populations in Sweden, the Netherlands, and the United States. *Journal of clinical microbiology*. 50 (11), 3548–3555.

Clark, J.L., Espinosa, B., Leon, S.R., Hall, E.R., *et al.* (2008) Absence of lymphogranuloma venereum infection among high-risk men who have sex with men in Lima, Peru. *International journal of STD & AIDS*. 19 (6), 427–428.

Clatts, M.C., Goldsamt, L.A. & Yi, H. (2005) An emerging HIV risk environment: a preliminary epidemiological profile of an MSM POZ Party in New York City. *Sexually transmitted infections*. 81 (5), 373–376.

Clinical Effectiveness Group of the British Association for Sexual, H. and & HIV (n.d.) *Management of LGV [2006]*.

Cohen, M.S. & Pilcher, C.D. (2005) Amplified HIV transmission and new approaches to HIV prevention. *The Journal of infectious diseases*. 191 (9), 1391–1393.

Cohen, S.E., Chew Ng, R.A., Katz, K.A., Bernstein, K.T., *et al.* (2011) Repeat syphilis among men who have sex with men in California, 2002-2006: implications for syphilis elimination efforts. *American journal of public health*. 102 (1), e1–8.

Cosentino, L.A., Campbell, T., Jett, A., Macio, I., *et al.* (2012) Use of nucleic acid amplification testing for diagnosis of anorectal sexually transmitted infections. *Journal of clinical microbiology*. 50 (6), 2005–2008.

Coutts, W.E. (1950) Lymphogranuloma venereum; a general review. *Bulletin of the World Health Organization*. 2 (4), 545–562.

Crepaz, N., Marks, G., Liau, A., Mullins, M.M., *et al.* (2009) Prevalence of unprotected anal intercourse among HIV-diagnosed MSM in the United States: a meta-analysis. *AIDS (London, England)*. 23 (13), 1617–1629.

Cullberg, M., Bratt, G., Petrersson, F. & Karlsson, A. (2009) [Rectal and pharyngeal Chlamydia common among men who have sex with men]. *Lakartidningen*. 106 (1-2), 24–27.

Daly, M., Hevey, D. & Regan, C. (2011) The role of perceived risk in general practitioners' decisions to inform partners of HIV-infected patients. *British Journal of Health Psychology*. 16273–287.

Danta, M., Brown, D., Bhagani, S., Pybus, O.G., *et al.* (2007) Recent epidemic of acute hepatitis C virus in HIV-positive men who have sex with men linked to high-risk sexual behaviours. *AIDS (London, England)*. 21 (8), 983–991.

Davies, P., Reid, D. & Weatherburn, P. (2002) *Putting it about: Health promotion for gay men with higher numbers of sexual partners*. . [Online]. 2002. Available from: http://www.sigmaresearch.org.uk/files/report2002b.pdf.

Davis, M. & Flowers, P. (2011) Love and HIV serodiscordance in gay men's accounts of life with their regular partners. *Culture Health & Sexuality*. 13 (7), 737–749.

Davis, T.W. & Goldstone, S.E. (2009) Sexually transmitted infections as a cause of proctitis in men who have sex with men. *Diseases of the colon and rectum*. 52 (3), 507–512.

De, P., Singh, A.E., Wong, T., Yacoub, W., *et al.* (2004) Sexual network analysis of a gonorrhoea outbreak. *Sexually Transmitted Infections*. 80 (4), 280–285.

Delpech, V. (2012) *Health System Concerns Related to TasP and Most At Risk Populations: Example from the United Kingdom*. . [Online]. 2012. IAPAC, plenary session. Available from: http://www.iapac.org/tasp_prep/presentations/TPSlon12_Plenary3_Delpech.pdf.

Dodds, C., Bourne, A. & Weait, M. (2009) Responses to criminal prosecutions for HIV transmission among gay men with HIV in England and Wales. *Reproductive health matters*. 17 (34), 135–145.

Dodds, C., Weatherburn, P., Bourne, A., Hammond, G., *et al.* (2009) *Sexually charged: The views of gay and bisexual men on criminal prosecutions for sexual HIV transmission*. . [Online]. 2009. Available from: http://www.sigmaresearch.org.uk/files/report2009a.pdf.

Dodds, J.P., Johnson, A.M., Parry, J. V & Mercey, D.E. (2007) A tale of three cities: persisting high HIV prevalence, risk behaviour and undiagnosed infection in community samples of men who have sex with men. *Sexually Transmitted Infections*. 83 (5), 392–396.

Doherty, I.A., Adimora, A.A., Muth, S.Q., Serre, M.L., *et al.* (2011) Comparison of sexual mixing patterns for syphilis in endemic and outbreak settings. *Sexually transmitted diseases*. 38 (5), 378–384.

Dougan, S., Evans, B.G. & Elford, J. (2007) Sexually transmitted infections in Western Europe among HIV-positive men who have sex with men. *Sexually Transmitted Diseases*. 34 (10), 783–790.

Drumright, L.N. & Frost, S.D.W. (2010) Rapid social network assessment for predicting HIV and STI risk among men attending bars and clubs in San Diego, California. *Sexually transmitted infections*. 86 Suppl 3iii17–23.

Dubois-Arber, F., Jeannin, A., Lociciro, S. & Balthasar, H. (2012) Risk reduction practices in men who have sex with men in Switzerland: serosorting, strategic positioning, and withdrawal before ejaculation. *Archives of sexual behavior*. 41 (5), 1263–1272.

Eaton, L.A., Kalichman, S.C., Cain, D.N., Cherry, C., *et al.* (2007) Serosorting sexual partners and risk for HIV among men who have sex with men. *American journal of preventive medicine*. 33 (6), 479–485.

Edwards-Jackson, N., Phanuphak, N., Van Tieu, H., Chomchey, N., *et al.* (2012) HIV serostatus disclosure is not associated with safer sexual behavior among HIV-positive men who have sex with men (MSM) and their partners at risk for infection in Bangkok, Thailand. *AIDS research and therapy*. 9 (1), 38.

Egger, M., Davey Smith, G. & Altman, D.G. (2001) 2nd Edition Book, Edited. *Systematic reviews in health care: meta-analysis in context*. London, BMJ Publishing Group.

Elford, J., Bolding, G. & Sherr, L. (2001) Seeking sex on the Internet and sexual risk behaviour among gay men using London gyms. *AIDS (Hagerstown)*. 15 (11), 1409–1415.

Elford, J., Ibrahim, F., Bukutu, C. & Anderson, J. (2007) Sexual behaviour of people living with HIV in London: implications for HIV transmission. *AIDS*. 21 Suppl 1S63–70.

Factor, R., Kawachi, I. & Williams, D.R. (2011) Understanding high-risk behavior among non-dominant minorities: a social resistance framework. *Social science & medicine (1982)*. 73 (9), 1292–1301.

Farhi, D., Gerhardt, P., Falissard, B., Poupet, H., *et al.* (2007) Increasing rates of quinolone-resistant Neisseria gonorrhoeae in Paris, France. *Journal of the European Academy of Dermatology and Venereology : JEADV*. 21 (6), 818–821.

Fenton, K.A. (2004) A multilevel approach to understanding the resurgence and evolution of infectious syphilis in Western Europe. *Euro Surveillance*. 9 (12), 3–4.

Fenton, K.A. & Imrie, J. (2005) Increasing rates of sexually transmitted diseases in homosexual men in Western europe and the United States: why? *Infectious disease clinics of North America*. 19 (2), 311–331.

Fenton, K.A., Mercer, C.H., Johnson, A.M., Byron, C.L., *et al.* (2005) Reported sexually transmitted disease clinic attendance and sexually transmitted infections in Britain: prevalence, risk factors, and proportionate population burden. *The Journal of infectious diseases*. 191 Suppl S127–38.

Flowers, P., Duncan, B. & Frankis, J. (2000) Community, responsibility and culpability: HIV risk-management amongst Scottish gay men. *Journal of Community & Applied Social Psychology*. 10 (4), 285–300.

Forrest, D.W., Metsch, L.R., LaLota, M., Cardenas, G., *et al.* (2010) Crystal Methamphetamine Use and Sexual Risk Behaviors among HIV-Positive and HIV-Negative Men Who Have Sex with Men in South Florida. *Journal of Urban Health-Bulletin of the New York Academy of Medicine*. 87 (3), 480–485.

Foss, A.M., Vickerman, P.T., Chalabi, Z., Mayaud, P., *et al.* (2009) Dynamic modeling of herpes simplex virus type-2 (HSV-2) transmission: issues in structural uncertainty. *Bulletin of mathematical biology*. 71 (3), 720–749.

French, P., Ison, C.A. & Macdonald, N. (2005) Lymphogranuloma venereum in the United Kingdom. *Sexually Transmitted Infections*. 81 (2), 97–98.

Frost, D.M., Stirratt, M.J. & Ouellette, S.C. (2008) Understanding why gay men seek HIV-seroconcordant partners: intimacy and risk reduction motivations. *Culture, health & sexuality*. 10 (5), 513–527.

Fung, M., Scott, K.C., Kent, C.K. & Klausner, J.D. (2007) Chlamydial and gonococcal reinfection among men: a systematic review of data to evaluate the need for retesting. *Sexually transmitted infections*. 83 (4), 304–309.

Garnett, G.P. & Anderson, R.M. (1994) Balancing sexual partnerships in an age and activity stratified model of HIV transmission in heterosexual populations. *IMA journal of mathematics applied in medicine and biology.* 11 (3), 161–192.

Garnett, G.P., Cousens, S., Hallett, T.B., Steketee, R., *et al.* (2011) Mathematical models in the evaluation of health programmes. *Lancet*. 378 (9790), 515–525.

Gebhardt, M. & Goldenberger, D. (2005) Lymphogranuloma venereum (LGV) serotype L2 in Switzerland, 2003-2005. *Euro Surveillance*. 10 (12).

Geisler, W.M., Morrison, S.G. & Bachmann, L.H. (2008) Absence of lymphogranuloma venereum strains among rectal Chlamydia trachomatis outer membrane protein A genotypes infecting women and men who have sex with men in Birmingham, Alabama. *Sexually Transmitted Diseases*. 35 (10), 856–858.

German, R.R., Lee, L.M., Horan, J.M., Milstein, R.L., *et al.* (2001) Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports / Centers for Disease Control*. 50 (RR-13), 1–35; quiz CE1–7.

Ghani, A.C. & Garnett, G.P. (2000) Risks of acquiring and transmitting sexually transmitted diseases in sexual partner networks. *Sexually Transmitted Diseases*. 27 (10), 579–587.

Ghosh, S.K. (2002) A study on the present scenario of STD management in an urban clinic in Kolkata. *Indian journal of dermatology, venereology and leprology*. 68 (2), 82–83.

Gilbart, V.L., Simms, I., Gobin, M., Oliver, I., *et al.* (2013) High-risk drug practices in men who have sex with men. *Lancet*. 381 (9875), 1358–1359.

Giraudon, I., Ruf, M., Maguire, H., Charlett, A., *et al.* (2008) Increase in diagnosed newly acquired hepatitis C in HIV-positive men who have sex with men across London and Brighton, 2002-2006: is this an outbreak? *Sexually transmitted infections*. 84 (2), 111–115.

Glick, S.N., Morris, M., Foxman, B., Aral, S.O., *et al.* (2012) A comparison of sexual behavior patterns among men who have sex with men and heterosexual men and women. *Journal of acquired immune deficiency syndromes (1999)*. 60 (1), 83–90.

Golden, M.R., Stekler, J., Hughes, J.P. & Wood, R.W. (2008) HIV serosorting in men who have sex with men: is it safe? *Journal of acquired immune deficiency syndromes (1999)*. 49 (2), 212–218.

Goldenberger, D., Dutly, F. & Gebhardt, M. (2006) Analysis of 721 Chlamydia trachomatis-positive urogenital specimens from men and women using lymphogranuloma venereum L2-specific real-time PCR assay. *Euro Surveillance*. 11 (10).

Gomez, G.B., Borquez, A., Caceres, C.F., Segura, E.R., *et al.* (2012) The Potential Impact of Pre-Exposure Prophylaxis for HIV Prevention among Men Who Have Sex with Men and Transwomen in Lima, Peru: A Mathematical Modelling Study. *PLoS medicine*. 9 (10).

Goodreau, S.M. & Golden, M.R. (2007) Biological and demographic causes of high HIV and sexually transmitted disease prevalence in men who have sex with men. *Sexually transmitted infections*. 83 (6), 458–462.

Goodreau, S.M., Golden, M.R., Goicochea, L.P. & Sanchez, J. (2007) Biological and demographic causes of high HIV and sexually transmitted disease prevalence in men who have sex with men Sexual role and transmission of HIV Type 1 among men who have sex with men, in Peru. *Sexually transmitted infections*. 83 (1368-4973 (Print)), 458–462.

Gorbach, Galea, Amani, Shin, *et al.* (2004) Don't ask, don't tell: patterns of HIV disclosure among HIV positive men who have sex with men with recent STI practising high risk behaviour in Los Angeles and Seattle. *Sexually Transmitted Infections*. 80 (6), 512–517.

Gorbach, P.M., Drumright, L.N., Daar, E.S. & Little, S.J. (2006) Transmission behaviors of recently HIV-infected men who have sex with men. *Jaids-Journal of Acquired Immune Deficiency Syndromes*. 42 (1), 80–85.

Gorbach, P.M., Weiss, R.E., Jeffries, R., Javanbakht, M., *et al.* (2011) Behaviors of recently HIV-infected men who have sex with men in the year postdiagnosis: effects of drug use and partner types. *Journal of acquired immune deficiency syndromes (1999)*. 56 (2), 176–182.

Götz, H., Nieuwenhuis, R., Ossewaarde, T., Bing Thio, H., *et al.* (2004) Preliminary report of an outbreak of lymphogranuloma venereum in homosexual men in the Netherlands, with implications for other countries in western Europe. *Euro Surveillance*. 8 (4).

Gotz, H.M., Van Doornum, G., Niesters, H.G., Den Hollander, J.G., *et al.* (2005) A cluster of acute hepatitis C virus infection among men who have sex with men--results from contact tracing and public health implications. *AIDS*. 19 (9), 969–974.

Grant, R.M., Lama, J.R., Anderson, P.L., McMahan, V., *et al.* (2010) Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *The New England journal of medicine*. 363 (27), 2587–2599.

Gray, R.H. & Wawer, M.J. (2008) Reassessing the hypothesis on STI control for HIV prevention. *The Lancet*. 371 (9630), 2064–2065.

Grov, C., Golub, S.A. & Parsons, J.T. (2010) HIV status differences in venues where highly sexually active gay and bisexual men meet sex partners: results from a pilot study. *AIDS Education and Prevention*. 22 (6), 496–508.

Gunn, R.A., Fitzgerald, S. & Aral, S.O. (2000) Sexually transmitted disease clinic clients at risk for subsequent gonorrhea and chlamydia infections: possible "core" transmitters. *Sexually transmitted diseases*. 27 (6), 343–349.

Guzman, R., Buchbinder, S., Mansergh, G., Vittinghoff, E., *et al.* (2006) Communication of HIV viral load to guide sexual risk decisions with serodiscordant partners among San Francisco men who have sex with men. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv*. 18 (8), 983–989.

Halioua, B., Bohbot, J.M., Monfort, L., Nassar, N., *et al.* (2006) Ano-rectal lymphogranuloma venereum: 22 cases reported in a sexually transmited infections center in Paris. *European journal of dermatology : EJD*. 16 (2), 177–180.

Halkitis, P.N., Green, K.A., Remien, R.H., Stirratt, M.J., *et al.* (2005) Seroconcordant sexual partnerings of HIV-seropositive men who have sex with men. *AIDS*. 19S77–S86.

Hallett, T.B., Smit, C., Garnett, G.P. & De Wolf, F. (2011) Estimating the risk of HIV transmission from homosexual men receiving treatment to their HIV-uninfected partners. *Sexually transmitted infections*. 87 (1), 17–21.

Halse, T.A., Musser, K.A. & Limberger, R.J. (2006) A multiplexed real-time PCR assay for rapid detection of Chlamydia trachomatis and identification of serovar L-2, the major cause of Lymphogranuloma venereum in New York. *Molecular and cellular probes*. 20 (5), 290–297.

Hamill, M., Benn, P., Carder, C., Copas, A., *et al.* (2007) The clinical manifestations of anorectal infection with lymphogranuloma venereum (LGV) versus non-LGV strains of Chlamydia trachomatis: a case-control study in homosexual men. *International journal of STD & AIDS*. 18 (7), 472–475.

Harawa, N.T., Williams, J.K., Ramamurthi, H.C. & Bingham, T.A. (2006) Perceptions towards condom use, sexual activity, and HIV disclosure among HIV-positive African American men who have sex with men: Implications for heterosexual transmission. *Journal of Urban Health-Bulletin of the New York Academy of Medicine*. 83 (4), 682–694.

Harris, S.R., Clarke, I.N., Seth-Smith, H.M.B., Solomon, A.W., *et al.* (2012) Whole-genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. *Nature Genetics*.

Hart, G.J. & Elford, J. (2010a) Sexual risk behaviour of men who have sex with men: emerging patterns and new challenges. *Current opinion in infectious diseases*. 23 (1), 39–44.

Hawker, J., Begg, N., Blair, I., Reintjes, R., *et al.* (2008) *Communicable Disease Control Handbook*. 2nd Ed. Blackwell Publishing Ltd.

Herida, M., Kreplack, G., Cardon, B., Desenclos, J.-C., *et al.* (2006) First case of urethritis due to Chlamydia trachomatis genovar L2b - Letter. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 43 (2), 268–269.

Herida, M., Sednaoui, P., Couturier, E., Neau, D., *et al.* (2005) Rectal lymphogranuloma venereum, France. *Emerging Infectious Diseases*. 11 (3), 505–506.

Herrmann, B., Eden, D., Hadad, R., Christerson, L., *et al.* (2012) Prevalence Trends of the New Variant of Chlamydia trachomatis in Four Counties of Sweden in 2007-2011. *Sexually transmitted diseases*. 39 (8), 648–650.

Hoff, C.C., Stall, R., Paul, J., Acree, M., *et al.* (1997) Differences in sexual behavior among HIV discordant and concordant gay men in primary relationships. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*. 14 (1), 72–78.

Hoff, Colleen, C., Chakravarty, Deepalika, *et al.* (2012) *Relationship Characteristics Associated with Sexual Risk Behavior Among MSM in Committed Relationships*. 26 (12), 738–745.

Von Holsterin, I., Fenton, A.F. & Ison, C. (2004) European network for surveillance of STIs (ESSTI) establishes working groups on lymphogranuloma venereum and HIV/STI prevention among MSM. *Euro Surveillance*. 8 (25).

Holt, R., Court, P., Vedhara, K., Nott, K.H., *et al.* (1998) The role of disclosure in coping with HIV infection. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv*. 10 (1), 49–60.

Holt, Rawstorne, Worth, Bittman, *et al.* (2011) Predictors of HIV Disclosure Among Untested, HIV-Negative and HIV-Positive Australian Men Who had Anal Intercourse with Their Most Recent Casual Male Sex Partner. *AIDS and Behavior*. 15 (6), 1128–1139.

Horvath, K.J., Oakes, J.M. & Rosser, B.R.S. (2008) Sexual negotiation and HIV serodisclosure among men who have sex with men with their online and offline partners. *Journal of Urban Health-Bulletin of the New York Academy of Medicine*. 85 (5), 744–758.

Hosmer, D. & Lemeshow, S. (2000) *Applied Logistic Regression*. 2nd edition. John Wiley & Sons, Ltd.

Houston, E., Sandfort, T., Dolezal, C. & Carballo-Dieguez, A. (2012) Depressive Symptoms Among MSM Who Engage in Bareback Sex: Does Mood Matter? *AIDS and Behavior*. 16 (8), 2209–2215.

Hu, F.B., Goldberg, J., Hedeker, D., Flay, B.R., *et al.* (1998) Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *American journal of epidemiology*. 147 (7), 694–703.

Hughes, G., Nichols, T., Peters, L., Bell, G., *et al.* (2012) Repeat infection with gonorrhoea in Sheffield, UK: predictable and preventable? *Sexually transmitted infections*.

Hughes, G., Paine, T. & Thomas, D. (2001) Surveillance of sexually transmitted infections in England and Wales. *Euro surveillance*. 6 (5), 71–80.

Ihekweazu, C., Maxwell, N., Organ, S. & Oliver, I. (2007) Is STI surveillance in England meeting the requirements of the 21st century? An evaluation of data from the South West Region. *Euro surveillance*. 12 (5).

Jaiswal, A.K., Banerjee, S., Matety, A.R. & Grover, S. (2002) Changing trends in sexually transmitted diseases in North Eastern India. *Indian journal of dermatology, venereology and leprology*. 68 (2), 65–66.

Jalal, H., Stephen, H., Alexander, S., Carne, C., *et al.* (2007) Development of real-time PCR assays for genotyping of Chlamydia trachomatis. *Journal of clinical microbiology*. 45 (8), 2649–2653.

Jebbari, H. (n.d.) *Public Health England - Protocol for Enhanced National Syphilis Surveillance*.[Online] Available from: www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1207554951587 [Accessed: 10 June 2010].

Jebbari, H., Alexander, S., Ward, H., Evans, B., *et al.* (2007) Update on lymphogranuloma venereum in the United Kingdom. *Sexually Transmitted Infections*. 83 (4), 324–326.

Jebbari, H., Simms, I., Conti, S., Marongiu, A., *et al.* (2011) Variations in the epidemiology of primary, secondary and early latent syphilis, England and Wales: 1999 to 2008. *Sexually Transmitted Infections*. 87 (3), 191–198.

Jeffrey, B.M., Suchland, R.J., Quinn, K.L., Davidson, J.R., *et al.* (2010) Genome sequencing of recent clinical Chlamydia trachomatis strains identifies loci associated with tissue tropism and regions of apparent recombination. *Infection and immunity*. 78 (6), 2544–2553.

Johnson, A.M., Mercer, C.H., Erens, B., Copas, A.J., *et al.* (2001) Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours. *Lancet*. 358 (9296), 1835–1842.

Johnson, L.F., Alkema, L. & Dorrington, R.E. (2010) A Bayesian approach to uncertainty analysis of sexually transmitted infection models. *Sexually Transmitted Infections*. 86 (3), 169–174.

Johnson, L.F., Coetzee, D.J. & Dorrington, R.E. (2005) Sentinel surveillance of sexually transmitted infections in South Africa: a review. *Sexually Transmitted Infections*. 81 (4), 287–293.

Johnson, M.O., Carrico, A.W., Chesney, M.A. & Morin, S.F. (2008) Internalized Heterosexism Among HIV-Positive, Gay-Identified Men: Implications for HIV Prevention and Care. *Journal of Consulting and Clinical Psychology*. 76 (5), 829–839.

Kaliaperumal, K. & Karthikeyan, K. (2008) Recent advances in management of genital ulcer disease and anogenital warts. *Dermatology and therapy.* 21 (3), 196–204.

Kanter, J., Koh, C., Razali, K., Tai, R., *et al.* (2011) Risk behaviour and HIV prevalence among men who have sex with men in a multiethnic society: a venue-based study in Kuala Lumpur, Malaysia. *International Journal of STD & AIDS*. 22 (1), 30–37.

Kapoor, S. (2008) Re-emergence of lymphogranuloma venereum. *Journal of the European Academy of Dermatology and Venereology : JEADV*. 22 (4), 409–416.

Keeling, M. & Rohani, P. (2008) *Modeling Infectious Diseases in Human and Animals*. New Jersey, Princeton University Press.

Kent, C.K., Chaw, J.K., Kohn, R.P., Chen, Y.Q., *et al.* (2004) Studies relying on passive retrospective cohorts developed from health services data provide biased estimates of incidence of sexually transmitted infections. *Sexually transmitted diseases*. 31 (10), 596–600.

Kirby, T. & Thornber-Dunwell, M. (2013) High-risk drug practices tighten grip on London gay scene. *The Lancet*. 381 (9861), 101–102.

Kirkwood, B.R. & Sterne, J.A.C. (2008) *Essential Medical Statistics*. 2nd Ed. Oxford, Balckwell Science.

Klint, M., Lofdahl, M., Ek, C., Airell, A., *et al.* (2006) Lymphogranuloma venereum prevalence in Sweden among men who have sex with men and characterization of Chlamydia trachomatis ompA genotypes. *Journal of clinical microbiology*. 44 (11), 4066–4071.

Klitzman, R., Exner, T., Correale, J., Kirshenbaum, S.B., *et al.* (2007) It's not just what you say: Relationships of HIV dislosure and risk reduction among MSM in the post-HAART era. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv*. 19 (6), 749–756.

Kretzschmar, M., Turner, K.M.E., Barton, P.M., Edmunds, W.J., *et al.* (2009) Predicting the population impact of chlamydia screening programmes: comparative mathematical modelling study. *Sexually transmitted infections*. 85 (5), 359–366.

Kropp, R.Y., Wong, T. & Canadian, L.G.V.W.G. (2005) Emergence of lymphogranuloma venereum in Canada. *Canadian Medical Association journal*. 172 (13), 1674–1676.

Kurtz, S.P., Buttram, M.E., Surratt, H.L. & Stall, R.D. (2012) Resilience, syndemic factors, and serosorting behaviors among HIV-positive and HIV-negative substance-using MSM. *AIDS education and prevention : official publication of the International Society for AIDS Education*. 24 (3), 193–205.

Van de Laar, M.J.W. (2006) The emergence of LGV in western Europe: what do we know, what can we do? *Euro Surveillance*. 11 (9), 146–148.

Van de Laar, M.J.W., Koedijk, F.D.H., Gotz, H.M. & De Vries, H.J.C. (2006) A slow epidemic of LGV in the Netherlands in 2004 and 2005. *Euro Surveillance*. 11 (9), 150–152.

Van de Laar, T., Pybus, O., Bruisten, S., Brown, D., *et al.* (2009) Evidence of a large, international network of HCV transmission in HIV-positive men who have sex with men. *Gastroenterology*. 136 (5), 1609–1617.

Lamontagne, S.D., Baster, K., Emmett, L., Nichols, T., *et al.* (2007) Incidence and reinfection rates of genital chlamydial infection among women aged 16-24 years attending general practice, family planning and genitourinary medicine clinics in England: a prospective cohort study by the Chlamydia Recall Study Advisory Group. *Sexually transmitted infections*. 83 (4), 292–303.

Larkins, S., Reback, C.J., Shoptaw, S. & Veniegas, R. (2005) Methamphetamine-dependent gay men's disclosure of their HIV status to sexual partners. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv*. 17 (4), 521–532.

Lattimore, S., Thornton, A., Delpech, V. & Elford, J. (2011a) Changing patterns of sexual risk behavior among London gay men: 1998-2008. *Sexually Transmitted Diseases*. 38 (3), 221–229.

Lee, S.S., Tam, D.K., Tan, Y., Mak, W.L., *et al.* (2009) An exploratory study on the social and genotypic clustering of HIV infection in men having sex with men. *AIDS (London, England)*. 23 (13), 1755–1764.

Leobon, A., Velter, A., Engler, K., Drouin, M.-C., *et al.* (2011) A relative profile of HIV-negative users of French websites for men seeking men and predictors of their regular risk taking: a comparison with HIV-positive users. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv*. 23 (1), 25–34.

Lewis, J.J., Donnelly, C.A., Mare, P., Mupambireyi, Z., *et al.* (2007) Evaluating the proximate determinants framework for HIV infection in rural Zimbabwe. *Sexually Transmitted Infections*. 83 Suppl 1i61–69.

Liang, K. & Zeger, S. (1986) Longitudinal data analysis using generalized linear models. *Biometrika*. 73 (1), 13–22.

Liassine, N., Caulfield, A., Ory, G., Restellini, A., *et al.* (2005) First confirmed case of lymphogranuloma venereum (LGV) in Switzerland. *Euro Surveillance*. 10 (7).

Liau, A., Millett, G. & Marks, G. (2006) Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. *Sexually transmitted diseases*. 33 (9), 576–584.

Lister, N.A., Tabrizi, S.N., Fairley, C.K., Smith, A., *et al.* (2004) Variability of the Chlamydia trachomatis omp1 gene detected in samples from men tested in male-only saunas in Melbourne, Australia. *Journal of clinical microbiology*. 42 (6), 2596–2601.

Mabey, D. & Peeling, R.W. (2002) Lymphogranuloma venereum. *Sexually Transmitted Infections*. 78 (2), 90–92.

Macdonald, N., Ison, C., Martin, I., Alexander, S., *et al.* (2005) Initial results of enhanced surveillance for lymphogranuloma venereum (LGV) in England. *Euro Surveillance*. 10 (1).

Mao, L., Kippax, S.C., Holt, M., Prestage, G.P., *et al.* (2011) Rates of condom and non-condom-based anal intercourse practices among homosexually active men in Australia: deliberate HIV risk reduction? *Sexually Transmitted Infections*. 87 (6), 489–493.

Marcus, U., Schmidt, A.J. & Hamouda, O. (2011) HIV serosorting among HIV-positive men who have sex with men is associated with increased self-reported incidence of bacterial sexually transmissible infections. *Sexual health*. 8 (2), 184–193.

Marcus, U., Zucs, P., Bremer, V., Hamouda, O., *et al.* (2004) Shigellosis - a re-emerging sexually transmitted infection: outbreak in men having sex with men in Berlin. *International journal of STD & AIDS*. 15 (8), 533–537.

Marks, G. & Crepaz, N. (2001) HIV-positive men's sexual practices in the context of self-disclosure of HIV status. *Journal of acquired immune deficiency syndromes*. 27 (1), 79–85.

Marks, G., Ruiz, M.S., Richardson, J.L., Reed, D., *et al.* (1994) Anal intercourse and disclosure of HIV-infection among seropositive gay and bisexual men. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*. 7 (8), 866–869.

McClelland, R.S., Lavreys, L., Katingima, C., Overbaugh, J., *et al.* (2005) Contribution of HIV-1 infection to acquisition of sexually transmitted disease: a 10-year prospective study. *The Journal of infectious diseases*. 191 (3), 333–338.

McConnell, Bragg, Shiboski, S., Grant, *et al.* (2010) Sexual seroadaptation: lessons for prevention and sex research from a cohort of HIV-positive men who have sex with men. *PLoS ONE*. 5 (1), e8831.

McFarland, W., Chen, Y.-H., Nguyen, B., Grasso, M., *et al.* (2012) Behavior, intention or chance? A longitudinal study of HIV seroadaptive behaviors, abstinence and condom use. *AIDS and behavior*. 16 (1), 121–131.

McFarland, W., Chen, Y.-H., Raymond, H.F., Nguyen, B., *et al.* (2011) HIV seroadaptation among individuals, within sexual dyads, and by sexual episodes, men who have sex with men, San Francisco, 2008. *AIDS care*. 23 (3), 261–268.

McLean, C.A., Stoner, B.P. & Workowski, K.A. (2007) Treatment of lymphogranuloma venereum. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America Infect Dis*. 44 Suppl 3S147–52.

McMillan, A., Kell, P., Ward, H., British Co-operative Clinical Group of British Association for Sexual, H., *et al.* (2008) Diagnosing chlamydia and managing proctitis in men who have sex with men: current UK practice. *Sexually Transmitted Infections*. 84 (2), 97–100.

Menon-Johansson, A.S., Young, Y., Maguire, H., Trail, P., *et al.* (2009) Surveillance of an infectious syphilis outbreak in south London, UK: a comparison of reporting systems. *International Journal of STD & AIDS*. 20 (3), 196–199.

Menza, T.W., Kerani, R.P., Handsfield, H.H. & Golden, M.R. (2011) Stable Sexual Risk Behavior in a Rapidly Changing Risk Environment: Findings from Population-Based Surveys of Men Who Have Sex with Men in Seattle, Washington, 2003-2006. *AIDS and Behavior*. 15 (2), 319–329.

Mercer, C.H., Fenton, K.A., Copas, A.J., Wellings, K., *et al.* (2004) Increasing prevalence of male homosexual partnerships and practices in Britain 1990-2000: evidence from national probability surveys. *AIDS (London, England)*. 18 (10), 1453–1458.

Michaels, S. & Lhomond, B. (2006) Conceptualization and measurement of homosexuality in sex surveys: a critical review. *Cadernos de Saúde Pública*. 22 (7), 1365–1374.

Miles, R.P.M. (1957) Rectal lymphogranuloma venereum. *British Journal of Surgery*. 45 (190), 180–188.

Millett, G.A., Flores, S.A., Peterson, J.L. & Bakeman, R. (2007) Explaining disparities in HIV infection among black and white men who have sex with men: a meta-analysis of HIV risk behaviors. *AIDS*. 21 (15), 2083–2091.

Mimiaga, M.J., Reisner, S.L., Bland, S., Cranston, K., *et al.* (2010) "It's a Quick Way to Get What You Want": A Formative Exploration of HIV Risk Among Urban Massachusetts Men Who Have Sex with Men Who Attend Sex Parties. *AIDS Patient Care and STDs*. 24 (10), 659–674.

Miner, M.H., Peterson, J.L., Welles, S.L., Jacoby, S.M., *et al.* (2009) How Do Social Norms Impact HIV Sexual Risk Behavior in HIV-Positive Men Who Have Sex with Men? Multiple Mediator Effects. *Journal of Health Psychology*. 14 (6), 761–770.

Morin, S.F., Shade, S.B., Steward, W.T., Carrico, A.W., *et al.* (2008) A behavioral intervention reduces HIV transmission risk by promoting sustained serosorting practices among HIV-infected men who have sex with men. *Journal of acquired immune deficiency syndromes (1999)*. 49 (5), 544–551.

Morre, S.A., Ouburg, S., Van Agtmael, M.A. & De Vries, H.J. (2008) Lymphogranuloma venereum diagnostics: from culture to real-time quadriplex polymerase chain reaction. *Sexually Transmitted Infections*. 84 (4), 252–253.

Morre, S.A., Spaargaren, J., Fennema, J.S., De Vries, H.J., *et al.* (2005) Real-time polymerase chain reaction to diagnose lymphogranuloma venereum. *Emerging Infectious Diseases*. 11 (8), 1311–1312.

Morris, M. (2010) Barking up the wrong evidence tree. Comment on Lurie & Rosenthal, "Concurrent partnerships as a driver of the HIV epidemic in sub-Saharan Africa? The evidence is limited". *AIDS and behavior*. 14 (1), 31–3; discussion 34–7.

Morton, A.N., Fairley, C.K., Zaia, A.M. & Chen, M.Y. (2006) Anorectal lymphogranuloma venereum in a Melbourne man. *Sexual Health*. 3 (3), 189–190.

Moskowitz, D.A. & Roloff, M.E. (2008) Vengeance, HIV disclosure, and perceived HIV transmission to others. *AIDS and Behavior*. 12 (5), 721–728.

Moskowitz, D.A. & Seal, D.W. (2011) Self-Esteem in HIV-Positive and HIV-Negative Gay and Bisexual Men: Implications for Risk-Taking Behaviors with Casual Sex Partners. *AIDS and Behavior*. 15 (3), 621–625.

Murphy, G., Charlett, A., Jordan, L.F., Osner, N., *et al.* (2004) HIV incidence appears constant in men who have sex with men despite widespread use of effective antiretroviral therapy. *AIDS (London, England)*. 18 (2), 265–272.

Murphy, R.D., Gorbach, P.M., Weiss, R.E., Hucks-Ortiz, C., *et al.* (2012) Seroadaptation in a Sample of Very Poor Los Angeles Area Men Who Have Sex with Men. *AIDS Behav*.

Nieuwenhuis, R.F., Ossewaarde, J.M., Van der Meijden, W.I. & Neumann, H.A.M. (2003) Unusual presentation of early lymphogranuloma venereum in an HIV-1 infected patient: effective treatment with 1 g azithromycin. *Sexually Transmitted Infections*. 79 (6), 453–455.

O'Byrne, P., Willmore, J., Bryan, A., Friedman, D.S., *et al.* (2013) Nondisclosure prosecutions and population health outcomes: examining HIV testing, HIV diagnoses, and the attitudes of men who have sex with men following nondisclosure prosecution media releases in Ottawa, Canada. *BMC public health*. 13 (1), 94.

O'Dell, B.L., Rosser, B.R.S., Miner, M.H. & Jacoby, S.M. (2008) HIV prevention altruism and sexual risk behavior in HIV-Positive men who have sex with men. *AIDS and Behavior*. 12 (5), 713–720.

O'Farrell, N., Morison, L., Moodley, P., Pillay, K., *et al.* (2008) Genital ulcers and concomitant complaints in men attending a sexually transmitted infections clinic: implications for sexually transmitted infections management. *Sexually Transmitted Diseases*. 35 (6), 545–549.

O'Leary, A., Fisher, H.H., Purcell, D.W., Spikes, P.S., *et al.* (2007) Correlates of risk patterns and Race/Ethnicity among HIV-Positive men who have sex with men. *AIDS and Behavior*. 11 (5), 706–715.

Office for National Statistics (n.d.) *The census in England and Wales*. [Online] 2010. Available from: http://www.ons.gov.uk/census/index.html

Overstreet, N.M., Earnshaw, V.A., Kalichman, S.C. & Quinn, D.M. (2012) Internalized stigma and HIV status disclosure among HIV-positive black men who have sex with men. *AIDS Care*.

Pantalone, D.W., Bimbi, D.S. & Parsons, J.T. (2008) Motivations for the recreational use of erectile enhancing medications in urban gay and bisexual men. *Sexually Transmitted Infections*. 84 (6), 458–462.

Parsons, J.T. & Bimbi, D.S. (2007) Intentional unprotected anal intercourse among sex who have sex with men: Barebacking-from behavior to identity. *AIDS and Behavior*. 11 (2), 277–287.

Parsons, J.T., Schrimshaw, E.W., Bimbi, D.S., Wolitski, R.J., *et al.* (2005) Consistent, inconsistent, and non-disclosure to casual sexual partners among HIV-seropositive gay and bisexual men. *AIDS (London, England)*. 19 Suppl 1S87–97.

Parsons, J.T., Schrimshaw, E.W., Wolitski, R.J., Halkitis, P.N., *et al.* (2005) Sexual harm reduction practices of HIV-seropositive gay and bisexual men: serosorting, strategic positioning, and withdrawal before ejaculation. *AIDS (London, England)*. 19 Suppl 1S13–25.

Parsons, J.T., Severino, J., Nanin, J., Punzalan, J.C., *et al.* (2006) Positive, negative, unknown: assumptions of HIV status among HIV-positive men who have sex with men. *AIDS education and prevention : official publication of the International Society for AIDS Education*. 18 (2), 139–149.

Patel, P., Taylor, M.M., Montoya, J.A., Hamburger, M.E., *et al.* (2006) Circuit parties: sexual behaviors and HIV disclosure practices among men who have sex with men at the White Party, Palm Springs, California, 2003. *AIDS care*. 18 (8), 1046–1049.

Pathela, P., Blank, S. & Schillinger, J.A. (2007) Lymphogranuloma venereum: old pathogen, new story. *Current Infectious Disease Reports*. 9 (2), 143–150.

Pelgrom, J.M., Vogelaers, D. & Colle, I. (2008) Hepatitis C-seroconversion within three to six months after having contracted clinical syphilis and/or lymphogranuloma venereum rectitis in five homosexually active, HIV seropositive men. *Acta clinica Belgica*. 63 (5), 335–338.

Pencheon, D., Guest, C., Melzer, D. & Gray, J.A.M. (2008) Oxford Handbook of Public Health Practice. *Oxford Handbooks*. p.691.

Peterman, T., Tian, L.H., Metcalf, C.A., Satterwhite, C.L., *et al.* (2006) High incidence of new sexually transmitted infections in the year following a sexually transmitted infection: a case for rescreening. *Annals of internal medicine*. 145 (8), 564–572.

Phillips, A.N., Cambiano, V., Nakagawa, F., Brown, A.E., *et al.* (2013) Increased HIV Incidence in Men Who Have Sex with Men Despite High Levels of ART-Induced Viral Suppression: Analysis of an Extensively Documented Epidemic Chiyu Zhang (ed.). *PLoS ONE*. 8 (2), e55312.

Phipps, W., Kent, C.K., Kohn, R. & Klausner, J.D. (2009) Risk factors for repeat syphilis in men who have sex with men, San Francisco. *Sexually Transmitted Diseases*. 36 (6), 331–335.

Pineiro, L., Montes, M., Gil-Setas, A., Camino, X., *et al.* (2009) Genotyping of Chlamydia trachomatis in an area of Northern Spain. *Enfermedades infecciosas y microbiologia clinica*.

Plettenberg, A., Von Krosigk, A., Stoehr, A. & Meyer, T. (2004) Four cases of lymphogranuloma venereum in Hamburg, 2003. *Euro Surveillance*. 8 (30).

Poppen, P.J., Relsen, C.A., Zea, M.C., Bianchi, F.T., *et al.* (2005) Serostatus disclosure, seroconcordance, partner relationship, and unprotected anal intercourse among HIV-positive Latino men who have sex with men. *AIDS Education and Prevention*. 17 (3), 227–237.

Poundstone, K.E., Strathdee, S.A. & Celentano, D.D. (2004) The social epidemiology of human immunodeficiency virus/acquired immunodeficiency syndrome. *Epidemiologic reviews*. 2622–35.

Public Health England (2011) *Genitourinary Medicine Clinic Activity Dataset (GUMCAD)*. 2011 (21 February,).

Public Health England (2010a) *Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP), Methods*. [Olnline] 2010. Available from: http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1195733849423

Public Health England (2010b) *LGV Outbreak Referral Form*. . [Online]. 2010. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1203348024386 [Accessed: 7 May 2012].

Public Health England (2010c) *Numbers accessing HIV care (SOPHID)*. . [Online]. 2010. Available from: http://www.hpa.org.uk/webw/HPAweb&Page&HPAwebAutoListDate/Page/1201094588844?p=1201094588844.

Public Health England (2010d) *STI Annual Data Tables*. . [Online]. 2010. Available from: http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1203348026613.

Punyacharoensin, N., Edmunds, W.J., De Angelis, D. & White, R.G. (2011) Mathematical models for the study of HIV spread and control amongst men who have sex with men. *European journal of epidemiology*. 26 (9), 695–709.

Purcell, D.W., Moss, S., Remien, R.H., Woods, W.J., *et al.* (2005) Illicit substance use, sexual risk, and HIV-positive gay and bisexual men: differences by serostatus of casual partners. *AIDS (London, England)*. 19 Suppl 1S37–47.

Quinn, T.C., Gaydos, C., Shepherd, M., Bobo, L., *et al.* (1996) Epidemiologic and microbiologic correlates of Chlamydia trachomatis infection in sexual partnerships. *JAMA : the journal of the American Medical Association*. 276 (21), 1737–1742.

Rawstorne, P., Fogarty, A., Crawford, J., Prestage, G., *et al.* (2007) Differences between HIV-positive gay men who “frequently”, “sometimes” or “never” engage in unprotected anal intercourse with serononconcordant casual partners: Positive Health cohort, Australia. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv*. 19 (4), 514–522.

Reidy, W.J. & Goodreau, S.M. (2010) The role of commercial sex venues in the HIV epidemic among men who have sex with men. *Epidemiology*. 21 (3), 349–359.

Rekart, M.L., Patrick, D.M., Chakraborty, B., Maginley, J.J.L., *et al.* (2003) Targeted mass treatment for syphilis with oral azithromycin. *Lancet*. 361 (9354), 313–314.

Relf, M. V, Bishop, T.L., Lachat, M.F., Schiavone, D.B., *et al.* (2009) A qualitative analysis of partner selection, HIV serostatus disclosure, and sexual behaviors among HIV-positive urban men. *AIDS Education and Prevention*. 21 (3), 280–297.

Rendina, H.J., Golub, S.A., Grov, C. & Parsons, J.T. (2012) Stigma and Sexual Compulsivity in a Community-Based Sample of HIV-Positive Gay and Bisexual Men. *AIDS and Behavior*. 16 (3), 741–750.

Richardson, D. & Goldmeier, D. (2007) Lymphogranuloma venereum: an emerging cause of proctitis in men who have sex with men. *International journal of STD & AIDS*. 18 (1), 11–4;quz15.

Richert, C.A., Peterman, T.A., Zaidi, A.A., Ransom, R.L., *et al.* (1993) A method for identifying persons at high risk for sexually transmitted infections: opportunity for targeting intervention. *American journal of public health*. 83 (4), 520–524.

Rietmeijer, C.A., Van Bemmelen, R., Judson, F.N. & Douglas, J.M. (2002) Incidence and repeat infection rates of Chlamydia trachomatis among male and female patients in an STD clinic: implications for screening and rescreening. *Sexually transmitted diseases*. 29 (2), 65–72.

Roest, R.W., Van der Meijden, W.I., European Branch of the International Union against Sexually Transmitted, I. & the European Office of the World Health, O. (2001) European guideline for the management of tropical genito-ulcerative diseases. *International journal of STD & AIDS*. 12 Suppl 378–83.

Rönn, M. (2009) *The emergence of Lymphogranuloma venereum in men who have sex with men*. Imperial College London.

Rönn, M. & Ward, H. (2011) The association between Lymphogranuloma venereum and HIV among men who have sex with men: systematic review and meta-analysis. *BMC infectious diseases*. 11:70.

Rothenberg, R.B., Potterat, J.J. & Woodhouse, D.E. (1996) Personal risk taking and the spread of disease: beyond core groups. *The Journal of infectious diseases*. 174 Suppl S144–9.

Rottingen, J.A., Cameron, D.W. & Garnett, G.P. (2001) A systematic review of the epidemiologic interactions between classic sexually transmitted diseases and HIV: how much really is known? *Sexually Transmitted Diseases*. 28 (10), 579–597.

Savage, E., Van de Laar, M.J., Gallay, A., Van der Sande, M., *et al.* (2009) Lymphogranuloma venereum in Europe, 2003-2008. *Euro Surveillance*. 14 (48), 10 December, 2009.

Saxon, C., Hughes, G., Ison, C. & UK LGV Case-Finding Study Group (2013) Increasing asymptomatic Lymphogranuloma venereum infection in the UK: results from a national case-finding study. In: *ISSTDR 2013 (abstract accepted)*. 2013

Schachter, J. & Moncada, J. (2005) Lymphogranuloma venereum: how to turn an endemic disease into an outbreak of a new disease? Start looking. *Sexually Transmitted Diseases*. 32 (6), 331–332.

Schachter, J. & Stephens, R.S. (2008) Biology of Chlamydia trachomatis. In: K K Holmes, P F Sparling, W E Stamm, P Piot, *et al.* (eds.). *Sexually Transmitted Diseases*. 4th edition. The McGraw-Hill Companies, Inc. pp. p. 555–574.

Sethi, G., Allason-Jones, E., Richens, J., Annan, N.T., *et al.* (2008) Lymphogranuloma venereum presenting as genital ulceration and inguinal syndrome in men who have sex with men in London, United Kingdom. *Sexually Transmitted Infections*. 85 (3), 165–170.

Sexually Transmitted Bacteria Reference Laboratory (n.d.) *Referral of specimend for confirmation of LGV*. . [Online]. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1194947393245 [Accessed: 7 May 2012].

Sfetcu, O., Irvine, N., Ngui, S.L., Emerson, C., *et al.* (2011) Hepatitis A outbreak predominantly affecting men who have sex with men in Northern Ireland, October 2008 to July 2009. *Euro surveillance*. 16 (9).

Sheon, N. & Crosby, G.M. (2004) Ambivalent tales of HIV disclosure in San Francisco. *Social Science & Medicine*. 58 (11), 2105–2118.

Sigma Research (2010) *Gay Men's Sex Survey*. 2011 (9 February).

Sigma Research (2009) *Vital Statistics 2008: The UK Gay Men's Sex Survey, London Strategic Health Authority (Aread sub-samples data report)*. p.27.

Simms, I., Fenton, K.A., Ashton, M., Turner, K.M., *et al.* (2005) The re-emergence of syphilis in the United Kingdom: the new epidemic phases. *Sexually Transmitted Diseases*. 32 (4), 220–226.

Simms, I., Macdonald, N., Ison, C., Martin, *et al.* (2004) Enhanced surveillance of lymphogranuloma venereum (LGV) begins in England. *Euro Surveillance*. 8 (41).

Sirivongrangson, P., Lolekha, R., Charoenwatanachokchai, A., Siangphoe, U., *et al.* (2012) HIV Risk Behavior Among HIV-Infected Men Who have Sex with Men in Bangkok, Thailand. *AIDS and Behavior*. 16 (3), 618–625.

Smelov, V., Morre, S.A., De Vries, H.J. & De Vries, H. (2008) Are serological chlamydia-specific markers useful to detect asymptomatic cases of lymphogranuloma venereum proctitis? *Sexually Transmitted Infections*. 84 (1), 77;authorrey77–8.

Van der Snoek, E.M., Ossewaarde, J.M., Van der Meijden, W.I., Mulder, P.G.H., *et al.* (2007a) Are serological chlamydia-specific markers useful to detect asymptomatic cases of lymphogranuloma venereum proctitis? - Authors' reply. *Sexually Transmitted Infections*. 84 (1), 77–78.

Van der Snoek, E.M., Ossewaarde, J.M., Van der Meijden, W.I., Mulder, P.G.H., *et al.* (2007b) The use of serological titres of IgA and IgG in (early) discrimination between rectal infection with non-lymphogranuloma venereum and lymphogranuloma venereum serovars of Chlamydia trachomatis. *Sexually Transmitted Infections*. 83 (4), 330–334.

Somboonna, N., Wan, R., Ojcius, D.M., Pettengill, M.A., *et al.* (2011) Hypervirulent Chlamydia trachomatis clinical strain is a recombinant between lymphogranuloma venereum (L(2)) and D lineages. *mBio*. 2 (3), e00045–11.

Spaargaren, J., Fennema, H.S.A., Morre, S.A., De Vries, H.J.C., *et al.* (2005) New lymphogranuloma venereum Chlamydia trachomatis variant, Amsterdam. *Emerging Infectious Diseases*. 11 (7), 1090–1092.

Spaargaren, J., Schachter, J., Moncada, J., De Vries, H.J.C., *et al.* (2005) Slow epidemic of lymphogranuloma venereum L2b strain. *Emerging Infectious Diseases*. 11 (11), 1787–1788.

Stamm, W.E. (2008a) Chlamydia trachomatis infections of the adult (Chapter 32). In: K K Holmes, P F Sparling, W E Stamm, P Piot, *et al.* (eds.). *Sexually Transmitted Diseases*. 4th edition. The McGraw-Hill Companies, Inc. pp. p. 575–593.

Stamm, W.E. (2008b) Lymphogranuloma venereum. In: K K Holmes, P F Sparling, W E Stamm, P Piot, *et al.* (eds.). *Sexually Transmitted Diseases*. 4th edition. The McGraw-Hill Companies, Inc. pp. p. 595–605.

STATA (n.d.) *Stata 12 help for tabulate twoway*. . [Online]. Available from: http://www.stata.com/help.cgi?tabulate+twoway [Accessed: 1 October 2012].

Sulkowski, M.S., Mast, E.E., Seeff, L.B. & Thomas, D.L. (2000) Hepatitis C virus infection as an opportunistic disease in persons infected with human immunodeficiency virus. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 30 Suppl 1S77–84.

Le Talec & Jablonsky (2008) *Seroadaptation instead of serosorting : a broader concept and a more precise process model - The WARNING*. . [Online]. 2008. Available from: http://www.thewarning.info/article.php?id_article=0249 [Accessed: 18 March 2013].

Theodore, P.S., Duran, R.E.F., Antoni, M.H. & Fernandez, M.I. (2004) Intimacy and sexual behavior among HIV-positive men-who-have-sex-with-men in primary relationships. *AIDS and Behavior*. 8 (3), 321–331.

Thomas, J.C. & Tucker, M.J. (1996) The development and use of the concept of a sexually transmitted disease core. *The Journal of infectious diseases*. 174 Suppl S134–43.

Thomson, N.R. (2008) Vive la difference. *Nature reviews. Microbiology*. 6 (7), 502–503.

Timen, A., Hulscher, M.E.J.L., Vos, D., Van de Laar, M.J.W., *et al.* (2008) Control measures used during lymphogranuloma venereum outbreak, Europe. *Emerging Infectious Diseases*. 14 (4), 573–578.

Tinmouth, J., Gilmour, M.W., Kovacs, C., Kropp, R., *et al.* (2008) Is there a reservoir of sub-clinical lymphogranuloma venereum and non-LGV Chlamydia trachomatis infection in men who have sex with men? *International journal of STD & AIDS*. 19 (12), 805–809.

Torrone, E.A., Satterwhite, C.L., Scholes, D., Yu, O., *et al.* (2013) Estimating chlamydia re-infection rates: an empirical example. *Sexually transmitted infections*. [Online] Available from: doi:10.1136/sextrans-2012-050970.

Truong, H.M., Truong, H.-H.M., Kellogg, T., Klausner, J.D., *et al.* (2006) Increases in sexually transmitted infections and sexual risk behaviour without a concurrent increase in HIV incidence

among men who have sex with men in San Francisco: a suggestion of HIV serosorting? *Sexually transmitted infections*. 82 (6), 461–466.

UNAIDS (2011) *UNAIDS Terminology Guidelines*. . [Online]. 2011. Available from: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2118_terminology-guidelines_en.pdf [Accessed: 25 September 2012].

Vall Mayans, M., Sanz Colomo, B. & Ossewaarde, J.M. (2005) First case of LGV confirmed in Barcelona. *Euro Surveillance*. 10 (2), E050203.2.

Vallabhaneni, S., Li, X., Vittinghoff, E., Donnell, D., *et al.* (2012) Seroadaptive Practices: Association with HIV Acquisition among HIV-Negative Men Who Have Sex with Men. *PloS one*. [Online] 7 (10), e45718. Available from: doi:10.1371/journal.pone.0045718.

Vallabhaneni, S., McConnell, J.J., Loeb, L., Hartogensis, W., *et al.* (2013) Changes in Seroadaptive Practices from before to after Diagnosis of Recent HIV Infection among Men Who Have Sex with Men. *PloS one*. 8 (2), e55397.

Vanable, P.A., Carey, M.P., Brown, J.L., Littlewood, R.A., *et al.* (2012) What HIV-Positive MSM Want from Sexual Risk Reduction Interventions: Findings from a Qualitative Study. *AIDS and Behavior*. 16 (3), 554–563.

Vanable, P.A., Ostrow, D.G. & McKirnan, D.J. (2003) Viral load and HIV treatment attitudes as correlates of sexual risk behavior among HIV-positive gay men. *Journal of Psychosomatic Research*. 54 (3), 263–269.

Vargas-Leguas, H., Garcia de Olalla, P., Arando, M., Armengol, P., *et al.* (2012) Lymphogranuloma venereum: a hidden emerging problem, Barcelona, 2011. *Euro surveillance*. 17 (2).

Velicko, I., Cullberg, M., Bratt, G., Mamlov, G., *et al.* (2009) [Lymphogranuloma venereum--increased spread in Sweden. A rare Chlamydia subtype which become more and more common among men who have sex with men]. *Lakartidningen*. 106 (1-2), 28–31.

Velter, A., Bouyssou-Michel, A., Arnaud, A. & Semaille, C. (2009) Do men who have sex with men use serosorting with casual partners in France? Results of a nationwide survey (ANRS-EN17-Presse Gay 2004). *Euro surveillance*. 14 (47).

Van de Ven, P., Kippax, S., Crawford, J., Rawstorne, P., *et al.* (2002) In a minority of gay men, sexual risk practice indicates strategic positioning for perceived risk reduction rather than unbridled sex. *AIDS care*. 14 (4), 471–480.

Van de Ven, P., Mao, L.M., Fogarty, A., Rawstorne, P., *et al.* (2005) Undetectable viral load is associated with sexual risk taking in HIV serodiscordant gay couples in Sydney. *AIDS*. 19 (2), 179–184.

Vermij, P. (2005) Scientists charged with choosing publication over public health. *Nature medicine.* 11 (5), 461.

Vickerman, P.T., Ndowa, F., O'Farrell, N., Steen, R., *et al.* (2009) Using mathematical modelling to estimate the impact of periodic presumptive treatment on the transmission of STIs and HIV amongst female sex workers. *Sexually Transmitted Infections*.

De Vries, H.J., Smelov, V., Middelburg, J.G., Pleijster, J., *et al.* (2009) Delayed microbial cure of lymphogranuloma venereum proctitis with doxycycline treatment. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 48 (5), e53–6.

De Vries, H.J.C., Van der Bij, A.K., Fennema, J.S.A., Smit, C., *et al.* (2008) Lymphogranuloma venereum proctitis in men who have sex with men is associated with anal enema use and high-risk behavior. *Sexually Transmitted Diseases*. 35 (2), 203–208.

De Vries, H.J.C., Van der Helm, J.J., Schim van der Loeff, M.F. & Van Dam, A.P. (2009) Multidrug-resistant Neisseria gonorrhoeae with reduced cefotaxime susceptibility is increasingly common in men who have sex with men, Amsterdam, the Netherlands. *Euro surveillance*. 14 (37).

De Vrieze, N.H.N., Van Rooijen, M., Van der Loeff, M.F.S. & De Vries, H.J.C. (2013) Anorectal and inguinal lymphogranuloma venereum among men who have sex with men in Amsterdam, the Netherlands: trends over time, symptomatology and concurrent infections. *Sexually transmitted infections*.

Waalboer, R., Van der Snoek, E.M., Van der Meijden, W.I., Mulder, P.G.H., *et al.* (2006) Analysis of rectal Chlamydia trachomatis serovar distribution including L2 (lymphogranuloma venereum) at the Erasmus MC STI clinic, Rotterdam. *Sexually Transmitted Infections*. 82 (3), 207–211.

Ward, H., Alexander, S., Carder, C., Dean, G., *et al.* (2009) The prevalence of lymphogranuloma venereum infection in men who have sex with men: results of a multicentre case finding study. *Sexually Transmitted Infections*. 85 (3), 173–175.

Ward, H., Macdonald, N., Ronn, M., Dean, G., *et al.* (2011) O1-S07.06 Lymphogranuloma venereum in the UK: is there evidence for rectal to rectal transmission? Results of a multicenter case control study. *Sexually Transmitted Infections*. 87 (Suppl 1), A40–A40.

Ward, H., Martin, I., Macdonald, N., Alexander, S., *et al.* (2007) Lymphogranuloma venereum in the United Kingdom. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 44 (1), 26–32.

Ward, H. & Rönn, M. (2010) Contribution of sexually transmitted infections to the sexual transmission of HIV. *Current opinion in HIV and AIDS*. 5 (4), 305–310.

Wasserheit, J.N. & Aral, S.O. (1996) The dynamic topology of sexually transmitted disease epidemics: implications for prevention strategies. *The Journal of infectious diseases*. 174 Suppl S201–13.

Watts, C., Zimmerman, C., Foss, A.M., Hossain, M., *et al.* (2010) Remodelling core group theory: the role of sustaining populations in HIV transmission. *Sexually transmitted infections*. 86 Suppl 3iii85–92.

Wei, C., Lim, S.H., Guadamuz, T.E. & Koe, S. (2012) HIV Disclosure and Sexual Transmission Behaviors among an Internet Sample of HIV-positive Men Who Have Sex with Men in Asia: Implications for Prevention with Positives. *AIDS and Behavior*. 16 (7), 1970–1978.

Wei, C., Raymond, H.F., Guadamuz, T.E., Stall, R., *et al.* (2011) Racial/Ethnic differences in seroadaptive and serodisclosure behaviors among men who have sex with men. *AIDS and behavior*. 15 (1), 22–29.

White, J. (2009) Manifestations and management of lymphogranuloma venereum. *Current opinion in infectious diseases*. (22), 57–66.

White, J. & Ison, C. (2008) Lymphogranuloma venereum: what does the clinician need to know? *Clinical medicine : journal of the Royal College of Physicians of London*. 8 (3), 327–330.

White, P.J., Ward, H., Cassell, J.A., Mercer, C.H., *et al.* (2005) Vicious and virtuous circles in the dynamics of infectious disease and the provision of health care: gonorrhea in Britain as an example. *The Journal of infectious diseases*. 192 (5), 824–836.

Whittington, W.L., Kent, C., Kissinger, P., Oh, M.K., *et al.* (2001) Determinants of persistent and recurrent Chlamydia trachomatis infection in young women: results of a multicenter cohort study. *Sexually transmitted diseases*. 28 (2), 117–123.

Williamson, L.M., Dodds, J.P., Mercey, D.E., Hart, G.J., *et al.* (2008) Sexual risk behaviour and knowledge of HIV status among community samples of gay men in the UK. *AIDS (London, England)*. 22 (9), 1063–1070.

Wilson, D.P., Regan, D.G., Heymer, K.-J., Jin, F., *et al.* (2010) Serosorting may increase the risk of HIV acquisition among men who have sex with men. *Sexually transmitted diseases*. 37 (1), 13–17.

Wolitski, R.J., Bailey, C.J., Leary, A.O., Gomez, C.A., *et al.* (2003) Self-perceived responsibility of HIV-Seropositive men who have sex with men for preventing HIV transmission. *AIDS and Behavior*. 7 (4), 363–372.

Wolitski, R.J., Gomez, C.A., Parsons, J.T. & Grp, S.S. (2005) Effects of a peer-led behavioral intervention to reduce HIV transmission and promote serostatus disclosure among HIV-seropositive gay and bisexual men. *AIDS*. 19S99–S109.

Xia, Q., Tholandi, M., Osmond, D.H., Pollack, L.M., *et al.* (2006) The effect of venue sampling on estimates of HIV prevalence and sexual risk behaviors in men who have sex with men. *Sexually Transmitted Diseases*. 33 (9), 545–550.

Zablotska, I.B., Imrie, J., Prestage, G., Crawford, J., *et al.* (2009) Gay men's current practice of HIV seroconcordant unprotected anal intercourse: serosorting or seroguessing? *AIDS care*. 21 (4), 501–510.

Zea, M.C., Reisen, C.A., Poppen, P.J. & Diaz, R.M. (2003) Asking and telling: Communication about HIV status among Latino HIV-positive gay men. *AIDS and Behavior*. 7 (2), 143–152.

Zeger, S.L. (1988) Commentary. *Statistics in Medicine*. 7 (1-2), 161–168.

Zou, H., Fairley, C.K., Guy, R., Bilardi, J., *et al.* (2013) Automated, computer generated reminders and increased detection of gonorrhoea, Chlamydia and syphilis in men who have sex with men. *PloS one*. 8 (4), e61972.