PHE HIV and STI data sharing policy

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. It does this through world-class science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. PHE is an operationally autonomous executive agency of the Department of Health.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Prepared by: Gwenda Hughes, Barry Evans, Fortune Ncube, Brian Rice, Sonia Ribeiro, Ian Simms, Geraldine Leong, Alison Brown, Emma Savage, Hamish Mohammed, Stephen Duffell.
For queries relating to this document, please contact: HIVSTIWebPortal@phe.gov.uk.

Any enquiries regarding this publication should be sent to us at publications@phe.gov.uk

Published April 2015
PHE publications gateway number: 2015022

This document is available in other formats on request. Please email
HIVSTIWebPortal@phe.gov.uk.

Contents

Document Control

Current version	PHE HIV and STI data sharing policy. _The details below only specify changes made on or after version 3.0._		
Current status	Final revision (4.1)		
Authors	Author name	Section amended/ added	Date amended
	Hamish Mohammed	V3.0 - Masking guidelines have been updated to take into account situations where data masking requirements have changed over time (due to revised data content). Also, an example (appendix 3) has been added.	01/10/2013
	Hamish Mohammed	V4.0- Updated section 3.1.3 to clarify that patient-level extracts can only be shared with PHE staff, including those holding visiting worker's agreements and honorary contracts. Also, further detail added to 'Summary Key Points' section.	12/03/2015
Reviewers	Reviewer name	Section reviewed	Date reviewed
	Gwenda Hughes, Holly Mitchell, Stephen Duffell, Sarah Woodhall, Vivian Hope	All	12/04/2015
	Hamish Mohammed	All	09/04/2015

Issued to	Person(s) issued to	Date issued
	Published on www.gov.uk	24/04/2015
File reference	www.gov.uk/government/publications/hiv-and-sti-data-sharing-policy	

Glossary

AIDS	Acquired Immunodeficiency Syndrome
CIDSC	Centre for Infectious Disease Surveillance and Control
CTAD	Chlamydia Testing Activity Dataset
FES	Field Epidemiology Services
FOI	Freedom of Information
GRASP	Gonoccocal Resistance to Antimicrobials Surveillance Programme
GUM	Genitourinary Medicine
GUMCADv2	Genitourinary Medicine Clinic Activity Dataset Version 2
HANDD	HIV and AIDS New Diagnoses and Deaths database
HIV	Human Immunodeficiency Virus
HPA	Health Protection Agency
ICH	Institute of Child Health
IDU	Injecting Drug User
LA	Local Authority (Lower-Tier)
LSOA	Lower Super Output Area
MSOA	Middle Super Output Area
NAISM	National Antenatal Infection Screening Monitoring
NCSP	National Chlamydia Screening Programme
NGO	Non-governmental Organisation
NHS	National Health Service
NRES	National Research Ethics Service
ONS	Office of National Statistics
PHE	Public Health England
PHE-C	Public Health England Centre
PHE-R	Public Health England Region
PCT	Primary Care Trust
PQ	Parliamentary Question
RITA	Recent HIV infection Testing Algorithm
SOP	Standard Operating Procedure
SOPHID	Survey of Prevalent HIV Infections Diagnosed
SHA	Strategic Health Authority
STI	Sexually Transmitted Infection
UA DBS	Unlinked Anonymous seroprevalence survey of HIV in neonatal dried blood spots
UTLA	Upper Tier Local Authority

Summary key points

- human immunodeficiency virus (HIV) and sexually transmitted infection (STI) data presented at the national, PHE region (PHE-R), or PHE centre (PHE-C) level may be published in hard copy and on the website
- data at the UTLA and/or LA level may only be published provided these have been assessed and considered not to be at risk of deductive disclosure
- the small cell size policy must be followed for all data that will be made publicly available. Suppression rules are to be applied to where cells are based on denominators less than 10,000 population. In this instance, cells with values from 1 to 4 inclusive must be anonymised and populated with '<5'. Where the anonymised cell could be deduced from other cells or the total, the next smallest cell size in the same row and/or column must also be anonymised and populated with '<x', such that 'x' is the value of the cell count rounded up to the nearest multiple of 10
- as a rule, when sharing data with small cell sizes (<5), a risk-assessment to prevent deductive disclosure and balance the public health benefits and risks to the individuals or PHE must be performed
- patient-level data extracts are not available to non-PHE staff
- any reproduction or analysis undertaken using data obtained from PHE must acknowledge the data source

1. Introduction

1.1. The HIV and STI Department (Centre for Infectious Disease Surveillance and Control, CIDSC) and the Small Cell Sizes subgroup of the Health Protection Agency (HPA, in 2013 the HPA was abolished and its functions were transferred to PHE) Caldicott Group have agreed the following policy with respect to storing, access, sharing and use of patient level and aggregated HIV and STI data. The policy is in line with the Caldicott principles and the HPA Caldicott Group consultation paper and Office for National Statistics (ONS) guidance. This policy will be reviewed annually and updated as required. This policy is designed to minimise the risk of deductive disclosure, however PHE are unable to guarantee that any such risk has been completely mitigated due to the availability of multiple other data sources.

1.2. This policy should be followed by all PHE staff, both at Colindale and in the regions, and all external users of HIV and STI data that have been collected and compiled by PHE. 'External users' refers to staff at LAs, service providers and anyone requesting data which are not routinely published. All users of HIV and STI data are expected to follow the same guidance given in this document with respect to publication. This is to ensure consistency between published statistics derived from the same source.

1.3. 'The data' refers to data from surveillance systems held at the HIV and STI Department, CIDSC, PHE, Colindale. These systems are described briefly in Appendix 1 and are as follows: the Genitourinary Medicine Clinic Activity Dataset (GUMCADv2), the Chlamydia Testing Activity Dataset (CTAD), the HIV and AIDS New Diagnoses and Deaths database, the Survey of Prevalent HIV Infections Diagnosed (SOPHID), the CD4 surveillance scheme, data from the Recent HIV infection Testing Algorithm (RITA) and the National Antenatal Infection Screening Monitoring (NAISM) programme.

2. Dealing with small cell sizes

2.1.   The small cell size policy must be followed for all data that will be made public or provided to those with no direct responsibility for infection control in the relevant population. ONS guidance states that where data tables become detailed, and counts in individual cells are small, the risk of identification increases and protection is needed[1]. HIV and STI data are considered in the high risk category.

2.2.   Geographic presentations

2.2.1. The lowest geography for data presentation in the public domain will be at the LA level, provided that these have been assessed and considered not to be at risk of deductive disclosure. STI and HIV surveillance data may be provided, by area of residence or area of reporting sites, where appropriate, as numbers or rates.

2.2.2. The number of diagnoses by lower level data (eg MSOA/LSOA level, ward level, clinic level) should not be released in the public domain because of the risk of deductive disclosure (please see section 4.2.2 for guidance on publishing rate data in maps).

2.3.   Disclosure by difference

2.3.1. This can occur when LA and UTLA data are presented side by side. The small differences in geographic boundaries can expose small pockets to disclosure. Such areas at risk of disclosure will be identified by PHE and the data masked as appropriate (see below).

2.4.   Masking cells that contain small numbers

2.4.1. All cells with values ranging from 1 to 4 are considered unsafe and care should be taken where a row or column is dominated by zeros[2]. When presenting data from denominators of less than 10,000 population, cells with values between 1 and 4 inclusive must be anonymised and populated with '<5' (see Example 1). This applies to data at UTLA, LA and lower geographies.

2.4.2. Where the anonymised cell could be deduced from other cells or the total, the next smallest cell size in the same row and/or column must also be anonymised and populated with '<x', such that 'x' is the value of the cell count rounded up to the nearest multiple of 10 (see Example 2).

2.4.3. Special attention should be paid to publishing data that are broken down by age-group, gender etc. Data can only be published without anonymisation if the denominator is greater or equal to 10,000 population. For this reason it is unlikely that publicly available data can be cross-tabulated at small geographic levels.

---

[1] ONS (2007) The review of the dissemination of health statistics in England, supporting paper.
[2] ONS (2006) Review of the dissemination of health statistics: Confidentiality Guidance

Example 1:

Local authority	Total	Male	Female	16 to 19	20 to 24
Population denominator	58,999	29,107	29,892	*2,956*	*3,238*
No of syphilis cases	4	3	1	0	3
Publish as	4	3	1	0	<5

Example 2:

	Male	Female	Total
Population denominator	11,000	*8,500*	19,500
No of syphilis cases	7	3	10
Publish as	<10	<5	10

2.4.4. Although the above rules are applicable in general, there may be situations where an additional risk assessment is warranted which may result in greater restrictions in data publication than those described above. For example:

- publication of data on STIs in children (people under 13 years of age) is particularly sensitive – diagnoses may be included in this age group for a number of reasons: it may be as a result of a recording error which led to the entry of an incorrect date of birth; it may involve the transmission of an infection that does not involve penetrative sex or other sexual activity; or it could involve child abuse – follow-up is not conducted to correct data entry errors and reporting errors so the numbers reported may not reflect the STIs that actually occur in this age group

- where data masking requirements have changed over time (due to revised data content) which may allow masked data to be deduced by comparing the content of previous/current data releases (see Appendix 3).

2.5. The operation of the ONS policy will be reviewed within PHE as required.

3. Levels of access to HIV and STI data

3.1. 'The data' can be accessed according to purpose, level of sharing (ie within a single or multiple LA/UTLAs), and applying the small cell size policy. The different levels of access are outlined below (see Table 1):

3.1.1. *High-level aggregated data:* These data are presented in tables, graphs or maps which summarise diagnosis numbers and rates by various patient characteristics at the national, PHE-R or PHE-C level. These data are usually made available publicly on the PHE website.

3.1.2. *Local-level aggregated data:* These data are presented in tables, graphs or maps which summarise diagnosis numbers and rates by patient characteristics at the UTLA, LA and lower levels, where appropriate (see section 4.2.2 for details on the sharing of MSOA/LSOA level data). These data are distributed, in confidence, to a range of stakeholders within the National Health Service (NHS), Department of Health/PHE and LAs for the purposes of planning and managing services, developing interventions and monitoring the effectiveness of health policies. Data at UTLA and LA levels may be published provided the small cell size policy (section 2) is followed at all times.

3.1.3. *Patient-level data extracts containing data on individual patient episodes*: Access to these data is restricted to nominated users who are accountable to PHE and must hold an honorary contract with PHE. Visiting students will only be able to access the data upon approval of a visiting worker's agreement and signature of the pertinent departmental data confidentiality and security agreement. All honorary contracts and visiting worker's agreements must be approved by the appropriate data custodian: the heads of STI and HIV surveillance at CIDSC or the regional epidemiologist or sexual health lead with responsibility for STIs and HIV. The guidelines around download and analysis of these extracts are presented in section 4.3.

3.1.4. *Use of HIV/STI public health surveillance data for research:* The primary purpose for the collection of these datasets is for public health monitoring, infection control and service provision. The data are not collected for primary research as defined by the National Research Ethics Service (NRES) and data at patient-level or lower geographies for research purposes only cannot currently be provided.

www.nres.nhs.uk/EasySiteWeb/GatewayLink.aspx?alId=355

Table 1: Levels of Access to HIV/STI data

Geography	PHE	Service provider[1]	UTLA/LA[1]	Other requesters[1] (eg researchers)
UTLA	+	+	+	+
LA	+	+	+	+
MSOA and lower[2]	+	-	+[3]	-
Patient level[2]	+	Own data only	-	-
Clinic level[2]	+	+	+	-

[1]Small cell size policy to be followed at all geography levels (section 2)
[2]MSOA/LSOA count data and patient or clinic level data are only available to those with an honorary contract or approved visiting worker's agreement with PHE
[3]See section 4.2.2 for detail on sharing of MSOA and LSOA rate data with LA/UTLA staff

3.2. Any data requests should be sent to the contact point for each surveillance system as listed in Appendix 2. Requesters will be required to complete a data request form detailing the data that are required and the purpose for which the data will be used. PHE staff dealing with data requests must follow the Standard Operating Procedure (SOP) for data requests.

3.3. All Freedom of Information (FOI) requests for HIV/STI data should be sent via the FOI office at PHE (foi@phe.gov.uk). The small cell size policy must be followed for all FOI requests.

3.4.   Press enquiries to CIDSC staff should be directed to the national PHE press office (infections-pressoffice@phe.gov.uk). Regional PHE staff should direct press queries either to the national PHE press office or the regional press office. The small cell size policy must be followed for all data requests including press queries.

3.5. Parliamentary Questions (PQs) should be sent via the PHE PQ office (pqs@phe.gov.uk). CIDSC staff must follow the SOP for PQs when dealing with a PQ. The small cell size policy also applies to data provided in response to PQs as these data are in the public domain.

4. Data sharing guidelines

The agreed policy for publishing 'the data' at the different levels is detailed in the following guidelines.

4.1.   Presentation and publication of high-level aggregated data

HIV, STI and National Antenatal Infections Screening Monitoring (NAISM) data tables at the level of PHE-C and above may be published by PHE, in hard copy or on the website. Anonymisation would not usually be required but may be considered appropriate in some cases. Any requests for aggregated NAISM data that are not within the public domain must be approved by the Unlinked Anonymous section head before publication.

4.2.   Presentation and publication of local-level aggregated data

4.2.1. HIV/STI data tables at the level of LA and/or UTLA may only be published, in hard copy or on the website, <u>provided small cell sizes are suitably anonymised</u> (see section 2 for details on the small cell size policy). Publishing proportions, percentages or rates should be the preferred method over publishing numbers. When in doubt, these tables should be reviewed on a case by case basis by the appropriate data custodian in consultation with CIDSC or the relevant regional Caldicott Guardian.

4.2.2. Maps of rates by UTLA/LA/MSOA/LSOA or tables of rates by UTLA/LA of residence, where rates are grouped into categories, may be published in public-facing reports. In the case of MSOA or LSOA-level data, tables with rates grouped into categories can be shared, in confidence, with non-PHE staff at LAs after a risk-assessment to prevent deductive disclosure and balance the public health benefits and risks to the individuals or PHE. The small cell size policy (see section 2) also applies to these categories ie the upper bound for the category with the lowest frequency must be greater than or equal to 5 (eg diagnosis rate per 100,000 population: 1 to 5, 6 to 9, 10 to 14, 15 to 19). MSOA or LSOA-level data tables cannot be published in hard copy of the website.

4.3.   Download and analysis of patient-level data extracts

4.3.1. Data extracts of patient-level information for epidemiological analysis must be downloaded by nominated users to a designated restricted-access folder on a network drive on a secure PHE server.

4.3.2. Data extracts should never be held on computer hard disks, laptops or any transportable storage media such as compact discs or memory sticks.

4.3.3. Data extracts must be deleted immediately after the purpose for which they were downloaded has been completed. Where extensive analyses are required, nominated users must request permission to keep data extracts beyond 1 year from the appropriate data custodian.

4.3.4. Data extracts must not be shared with non-nominated users. If complex statistical analysis of the data by non-nominated users is required, patient identifiers (patient clinic numbers, soundex, date of birth, postcode of residence)

must be removed from the data file and replaced with dummy numbers prior to analysis by the elected non-nominated user. Age and LSOA can be provided in lieu of date of birth and postcode, respectively. Such stripped down files must be managed according to principles 4.3.1 to 4.3.3 above.

4.4.   Acknowledging data sources

4.4.1. For data obtained from the HIV and STI Department to be published (either as provided or following additional analyses) in reports, peer-reviewed journals or on a website PHE must be acknowledged as follows: 'Data from xxx surveillance, Public Health England' (such that 'xxx' is GUMCADv2, NAISM, etc).

4.4.2. Prior approval is required from PHE for conducting analyses on data provided (this ensures the necessary caveats are considered).

4.4.3. Data from individual clinics must not be published without the prior consent and collaboration of the lead consultants at those clinics and PHE.

5. Responsibilities for adhering to policy

5.1. Use of 'the data': The data sharing policy permits the use of 'the data' by the 'data requester' within their organisation. Any data published by the requester should acknowledge PHE (see 4.4.1.).

5.2. Any breaches of this policy should be reported immediately, as follows:

5.2.1. Serious breaches, such as loss of patient-level data due to storage on non-permitted media (eg CDs, memory sticks etc.), should be reported to the relevant PHE institution (ie Colindale or region) and hospital, or community Caldicott and/or Associate Caldicott Guardians. Reports of serious breaches should be copied to all CIDSC and regional Associate Caldicott Guardians and to the HIV or STI Surveillance leads at CIDSC.

5.2.2. Breaches involving the publication of tabular aggregated data at UTLA level or below which have not been anonymised should be reported to the appropriate CIDSC or regional Associate Caldicott Guardian and copied to the STI or HIV Surveillance leads at CIDSC.

5.2.3. Minor breaches such as the publication of anonymised tabular aggregated data at LA level or below without permission from the relevant data custodian (ie the Heads of STI and HIV Surveillance at CIDSC Colindale or the Regional Epidemiologist or Sexual Health Lead with responsibility for STIs and HIV) should be reported to the appropriate CIDSC or regional Associate Caldicott Guardian and copied to the appropriate Surveillance lead at CIDSC.

6. Appendices

Appendix 1: Surveillance systems

Background to the surveillance systems covered by the policy is presented in sections A.1.1 to A.1.7. Terms relating to other surveillance systems in the HIV and STI department are in sections A.1.8. to A.1.11.

A.1.1 Genitourinary Medicine Clinic Activity Dataset Version 2

A.1.1.1. The Genitourinary Medicine Clinic Activity Dataset Version 2 (GUMCADv2) is used to monitor trends in new diagnoses of STIs and other sexual health problems and to determine which specific groups are at particular risk.

This information is used to inform the public health response by:
- improving the planning and management of services
- developing, adapting and refining interventions
- monitoring the effectiveness of sexual health policies

A.1.1.2. The GUMCADv2 return includes patient demographic details collected at patient registration at their first attendance at a GUM clinic and clinical and risk factor data collected during the patient consultation. The data are pseudo-anonymised (ie they contain the patient's clinic/hospital number but they do not contain patient-identifiable information such as name, date of birth, or postcode). These are sensitive patient-level data and their storage and access are under strict control.

A.1.2. Chlamydia Testing Activity Dataset

A.1.2.1. The Chlamydia Testing Activity Dataset (CTAD) is a universal disaggregate dataset for the collection of data on all NHS and NHS-commissioned chlamydia testing carried out in England. It replaced the National Chlamydia Screening Programme (NCSP) core data return and the non-NCSP non-GUM aggregate data return.

The data are used to inform the public health response by:

- improving the management and testing of chlamydia testing services

- monitoring the effectiveness of sexual health policies

A.1.2.2. The CTAD return includes patient demographic details collected at the time of screening. No names will be sent, but confidential data, including the patient's date of birth, postcode, clinic number and NHS number, will be collected. These are sensitive patient-level data and their storage and access are under strict control.

A.1.3. HIV and AIDS New Diagnoses and Deaths Database

A.1.3.1. The HIV and AIDS new diagnoses and deaths database (new HIV diagnoses) collects information from voluntary laboratory and clinician reports made in England, Wales and Northern Ireland. The information collected at diagnosis is on new diagnoses of HIV, first AIDS diagnosis and deaths in HIV-infected individuals aged 15 years and over. Scottish data (Health Protection Scotland) and data concerning paediatric infections (University College London Institute of Child Health) are collected separately and collated at PHE. Deaths in HIV-infected individuals are reported through voluntary clinician reports alongside matching-to-death records held by the Office for National Statistics.

This information is used to inform the public health response by:

- improving the planning and management of HIV testing and services
- developing, adapting and refining interventions
- monitoring the effectiveness of sexual health policies

A.1.3.2. Epidemiological information is collected to describe the characteristics of those newly diagnosed with HIV. The data are part-pseudo-anonymised ie they contain the patient's Soundex code (4-character code of surname), date of birth and clinic number but do not contain their name or residence postcode. They are highly sensitive patient-level data and their storage and access are under strict control.

A.1.4. Survey of prevalent HIV infections diagnosed

A.1.4.1. The survey of prevalent HIV infections diagnosed (SOPHID) collects annual data on all persons who attend for HIV-related care at an NHS site in England, Wales and Northern Ireland. Data from London (and parts of the South East) are collected twice yearly. Scottish data (Health Protection Scotland) and data on paediatric HIV infections (University College London Institute of Child Health) are collected separately and collated at PHE.

The data are used to inform the public health response by:

- estimating and characterising the number of people living with HIV infection in the UK
- calculating diagnosed HIV prevalence and estimating overall HIV prevalence
- improving the commissioning, planning and management of HIV testing and services
- developing, adapting and evaluating interventions
- developing clinical outcomes to audit the quality of services
- monitoring the effectiveness of sexual health policies

A.1.4.2. Epidemiological information is collected to describe the characteristics of diagnosed HIV-infected individuals accessing care and treatment. The data are part-pseudo-anonymised ie they contain the patient's Soundex code (4-character code of surname), date of birth, postcode and clinic number but does not contain their name.

They are highly sensitive patient-level data and their storage and access are under strict control.

A.1.5. CD4 Surveillance

A.1.5.1. The CD4 surveillance scheme monitors national trends in immuno-suppression among diagnosed HIV-infected adults. This is performed by analysis of CD4 cell counts. Data on patient's CD4 counts are collected from over 60 laboratories in England, Wales and Northern Ireland. Scottish data (Health Protection Scotland) are collected separately and collated at CIDSC.

Together with data from the HIV and AIDS New Diagnoses and Deaths database and SOPHID the data are used to inform the public health response by:
- monitoring levels of immuno-suppression among patients living with diagnosed HIV infection in the UK, by risk group and over time
- monitoring the proportion of newly HIV diagnosed patients who present at a late stage of infection (defined as <350 cells)
- monitoring the population effect of antiretroviral therapy

A.1.5.2. The data are part-pseudo-anonymised ie they contain the patient's Soundex code (4-character code of surname), date of birth, sex and laboratory number, but does not contain their name or residence postcode. They are highly sensitive patient-level data and their storage and access are under strict control.

A.1.6. The Recent Infection Testing Algorithm

A.1.6.1 The recent infection testing algorithm (RITA) surveillance system monitors the number of recently acquired HIV infections (in the last 6 months) among individuals newly diagnosed with HIV. Blood specimens are sent by laboratories in England and Northern Ireland to the Virus Reference Department at CIDSC where testing for recent infection is undertaken using the AxSym Avidity test. Results of the test are linked to the New HIV Diagnoses database.

Data from the RITA surveillance system are used to inform public health response by:
- monitoring the number of newly acquired HIV infections in England and Northern Ireland, over time and by risk group
- estimating HIV incidence in the UK
- identifying groups at most risk of acquiring HIV
- evaluating the impact of HIV prevention measures
- prioritising contact tracing

A.1.6.2 The data are pseudo-anonymised ie they contain the patient's Soundex code (4-character code of surname), date of birth, sex and laboratory number but does not contain their name or residence postcode. They are highly sensitive patient-level data and their storage and access are under strict control.

A.1.7. PHE National Antenatal Infection Screening Monitoring (NAISM) programme

A.1.7.1. The PHE national antenatal infection screening monitoring (NAISM) programme was established in 2004 to monitor the uptake and test results of antenatal screening at the national level.

A.1.7.2. Information on the uptake and results of antenatal screening for hepatitis B, HIV, syphilis and rubella susceptibility are collected at maternity units or trust level and supplied to the regional Field Epidemiology Services (FES) office where data are cleaned. Data are then collated and analysed at CIDSC.

The data are used to inform the public health response by:

- informing public health, clinical professionals and policy makers on the distribution and trends at the national and regional level
- detecting localised or national increases in hepatitis B, HIV or syphilis infection (including the detection of outbreaks) or susceptibility to rubella infection due to the failure of vaccination or other intervention programmes

A.1.8. Gonococcal resistance to antimicrobials surveillance programme (GRASP)

A.1.8.1. GRASP reports on a national and regional level in an annual report which is publicly available. Antibiotic resistance and epidemiological data are reported as percentages. Clinic data are only made available to the clinic providing the data.

A.1.9. Unlinked Anonymous seroprevalence survey of HIV in neonatal dried blood spots (UA DBS)

A.1.9.1. Deductive disclosure

No patient identifiers are contained within the dataset and the number of samples tested in the survey each year (~450,000) means that the issue of small cell sizes is rarely an issue.

A.1.9.2. Local level aggregated data

Residence based prevalence tables at UTLA and PHE-C levels are available on the website and sent to collaborators, regional PHE colleagues and other interested parties eg midwives, non-governmental organisations (NGOs). These tables can also be requested from regional sexual health leads. Annual supplementary datasets containing breakdowns by country of birth and age for England, London and outside London are also made available via the website and to the same mailing list as previously stated.

A.1.9.3. Ad hoc aggregate data

The HIV and STI policy document are followed replacing the confidentiality issues with deductive disclosure. The protocol within the systems ethics approval states that data cannot be presented where the denominator is <20. This cannot be changed without submitting an amendment to our ethics approval.

A.1.9.4. Patient level (disaggregate) data

Disaggregate level data is never shared for reasons of deductive disclosure. This is also stipulated in our ethics approval.

A.1.9.5. Storage of and access to patient level (disaggregate) data

Disaggregate data is stored securely in restricted folders accessed only by UA DBS staff and data transfer is via CyberArk.

A.1.9.6. Patient level (disaggregate) data analysis

Data analysis of disaggregate data is only undertaken by the principal investigator and their team or the senior investigator (based in the Institute for Child Health, ICH) and associated personnel.

A.1.10. Unlinked anonymous monitoring survey of people who inject drugs

A.1.10.1. The unlinked anonymous monitoring survey of people who inject drugs is an annual survey of injecting drug users (IDU) in contact with a reflective sample of specialist services for drug users. It provides data on the prevalence of HIV, hepatitis B and C and subject-reported data on key risk and protective behaviours among injecting drug users. Data from the survey are only available for England, Wales and Northern Ireland and by English region. Data tables, updated annually in July, are available on the PHE website.

A.1.11. Occupational Surveillance

A.1.11.1. The occupational surveillance scheme publishes aggregate data, by year and type of exposure, virus exposed to, occupation, location, etc. The number of reports received is presented at PHE-C level. The hepatitis C seroconversion cases are individually listed and detailed, but there is no overt information which would lead to the identification of the healthcare workers involved.

Appendix 2: Contacts for HIV/STI data

GUMCADv2: Data request form available at:

www.gov.uk/government/publications/gumcadv2-data-request-form

National Chlamydia Screening Programme: Please contact ncspdata@phe.gov.uk

New HIV Diagnoses

Survey of prevalent HIV infections diagnosed (SOPHID)

CD4 Surveillance

Recent HIV Infection Testing Algorithm

Data request form available at:

www.gov.uk/government/publications/hiv-and-aids-reporting-section-hars-data-request-form

Appendix 3: Example of revised masking requirements

This applies where data masking requirements have changed over time (due to revised data content) which may allow masked data to be deduced by comparing the content of previous/current data releases.

Table 1a: Data masking - previous release

STI	Gender	2008	2009	2010	2011	2012	Total
Syphilis	Male	9	8	*	*	*	34
	Female	6	5	*	*	*	18
	Total	15	13	12	6	6	52

Table 1b: Data masking revealed - previous release

STI	Gender	2008	2009	2010	2011	2012	Total
Syphilis	Male	9	8	* (7)	* (5)	* (5)	34
	Female	6	5	* (5)	* (1)	* (1)	18
	Total	15	13	12	6	6	52

Table 2a: Data masking - current release

STI	Gender	2008	2009	2010	2011	2012	Total
Syphilis	Male	9	8	7	*	*	34
	Female	6	5	5	*	*	20
	Total	15	13	12	6	8	54

Table 2b: Data masking revealed - current release

STI	Gender	2008	2009	2010	2011	2012	Total
Syphilis	Male	9	8	7	* (5)	* (5)	34
	Female	6	5	5	* (1)	* (3)	20
	Total	15	13	12	6	8	54

1/ The current release (Table 2a - circled in blue) presents data content that has been revised since the previous release (Table 1a). This results in a change to data masking requirements.

2/ The current release (Table 2a - circled in green) reveals data that was masked in the previous release (Table 1a)

3/ The data revealed in the current release (Table 2a - circled in green) can be used to unmask data from the previous release (Table 1b - circled in red)

4/ The data unmasked in the previous release (Table 1b - circled in red) can be used to unmask data in the current release (Table 2b - circled in red)