

Health and care 2006

# Review of the Dissemination of Health Statistics: Confidentiality Guidance

Working Paper 1: Confidentiality Protection – Legal and Policy Considerations

Office for National Statistics

1 Drummond Gate

London SW1V 2QQ

Tel: 020 7533 9233

Fax: 020 7533 9292

#### © Crown copyright 2006

Published with the permission of the Controller of Her Majesty's Stationery Office (HMSO)

You may re-use this publication (excluding logos) free of charge, in any format for research or private study or internal circulation within an organisation. You must re-use it accurately and not use it in a misleading context. The material must be acknowledged as crown copyright and you must give the title of the source publication. Where we have identified any third party copyright material you will need to obtain permission from the copyright holders concerned.

For any other use of this material please apply for a free Click-Use Licence for core material at: www.opsi.gov.uk/click-use

or write to: Office for Public Sector Information, St Clements House, 2-16 Colegate, Norwich NR3 1BQ.

Fax: 01603 723000

Email: hmsolicensing@cabinet-office.x.gsi.gov.uk

#### **Contact points**

For enquiries about this publication contact:

Statistical disclosure centre on 0845 601 3034

Email: info@statistics.gsi.gov.uk

For general enquiries, contact the National Statistics Customer Contact Centre:

Tel: 0845 601 3034

Minicom: 01633 812399

Email: info@statistics.gsi.gov.uk

Fax: 01633 652747

Post: Room 1015, Government Buildings, Cardiff Road, Newport NP10 8XG

#### **About the Office for National Statistics**

The Office for National Statistics (ONS) is the government agency responsible for compiling, analysing and disseminating economic, social and demographic statistics about the United Kingdom. It also administers the statutory registration of births, marriages and deaths in England and Wales. The Director of ONS is also the National Statistician and the Registrar General for England and Wales.

#### A National Statistics publication

National Statistics are produced to professional standards set out in the National Statistics Code of Practice. They are produced free from any political influence.

## Contents

1	Introduction	1
2	Why is confidentiality protection needed?	2
	Confidentiality protection in statistical principles – general considerations	.2
	Confidentiality protection in law – general considerations	.3
	General considerations of National Statistics	4
3	More on information laws	. 7
	Sources of confidentiality rights and protections for individuals	7
	Obligations to prohibit the disclosure of information about	
	individuals, and the authority to supply or acquire individual	
	information for statistics	9
	Rights of access to information used to produce statistics	10
Re	eferences	13



### 1 Introduction

This working paper provides more information on the legal and policy aspects of confidentiality protection to support the guidance for the review of the dissemination of health statistics.

## 2 Why is confidentiality protection needed?

Confidentiality protection in statistical principles – general considerations

It is a fundamental principle that the statistical records of individual persons, businesses or events used to produce Official Statistics are strictly confidential, and are to be used only for statistical purposes:

Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes. (UNECE 1992)

It is recognised that the quality of Official Statistics depends largely on the trust and cooperation of citizens. Trust and cooperation are gained when the privacy of individuals' information is respected. When a published statistic has adequate confidentiality protection the individual statistical records used to produce the statistic are not identifiable. This ensures the individual records are kept confidential to the producer of the statistic. A published statistic that keeps the individual records confidential ensures the individual records can not be used for unauthorised purposes. A published statistic should not allow any action or decision to be made relating to a specific individual.

For health statistics, it is essential that any published statistic respects the privacy of the information shared by individuals with health and statistics professionals. Failure to respect this privacy might result in harm or distress to a specific individual. Such a breach may have a wider effect. A breach of privacy in Official Statistics could damage the relationship of trust between private individuals and health and statistics professionals. Thus the public interest is served when statistical records are kept strictly confidential.

If this was the only public interest to be considered, the production of Official Statistics would be relatively easy. The challenge is to balance the public interest of confidentiality with the equally important public interest in the ready availability of statistical information. This is recognised in the UN Fundamental Principles:

1. Official statistics provide an indispensable element in the information system of a democratic society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information. (UNECE 1992)

Therefore confidentiality protection must be adequate to address the public interest in confidentiality of individual information, but not excessive so as to harm the practical utility of Official Statistics.

2

#### Confidentiality protection in law – general considerations

The legal framework for using health information that relates to identifiable individuals is complex. When adequate confidentiality protection is applied to this information the legal framework is dramatically simplified. To benefit from the simplification, the adequacy of the confidentiality protection should be such that the information is no longer likely to allow for the identification of specific individuals. Information that does not allow for the identification of individuals is not personal data and its disclosure to other health professionals or the public will not constitute a breach of data protection principles. Such information does not relate to an identifiable individual, and therefore its disclosure would not give cause for an action of breach of confidence. Information with adequate confidentiality protection does not relate to an individual's privacy or family, and therefore its disclosure would not be a breach of Article 8 of the Human Rights Act.

However, information that does allow individuals to be identified is likely to be subject to a duty of confidence, the Data Protection Act, the Human Rights Act, and any relevant prohibitions on disclosure in statute. Disclosure of information that allows for the identification of individuals may be in breach of one or more of these sources of information law.

Therefore the critical matter for producers of health statistics is to establish the necessary degree of confidentiality protection to ensure the release is of non-disclosive statistics only. Methods are available to assess the risk of a statistic allowing the identification of individuals, and to protect the statistics in such a way as to minimise the risk of a breach of information law.

The compilation of certain statistics may require the exchange of information that has not had confidentiality protection. Statistics for small areas, statistics that are derived from longitudinal information, and statistics that are derived from the linking or matching of information about individuals from difference sources, may require the exchange of information that allows for the identification of individuals. Applying confidentiality protection to information exchanged for these statistical purposes would limit or prevent the production of statistics of practical utility. Producers of such statistics need to be familiar with the legal framework for processing personal health information. Comprehensive legal guidance for the processing of individual's health information is available from the Information Commissioner. A guide to sharing data for statistical purposes, the ONS publication, *Data Sharing for a Statistical Purpose: A Practitioners' Guide to the Legal Framework* is available on the National Statistics website.

The compilation of patient-identifiable data for research should be conducted within the Research Governance Framework (Department of Health 2005). The statistical outputs of such research should reference the standards for confidentiality protection as set out in this guidance.

Some information relating to confidentiality rights, and the effect of the Freedom of Information Act, is in chapter 3. This information is not of direct

relevance to this guidance on confidentiality protection standards, but may be useful background for some users.

#### General considerations of National Statistics

The National Statistics Code of Practice states:

National Statistics will be produced in the interests of all citizens by protecting confidentiality.

The Code guarantees confidentiality:

Where data are collected or used for statistical purposes, we guarantee to protect confidentiality.

The Protocol on Data Access and Confidentiality requires (in part 1(d)) that for any National Statistic:

...statistical disclosure control measures will be adequate to ensure the confidentiality guarantee, and beyond that, as comprehensive as can be achieved without unduly compromising relevance, integrity and quality.

Thus the Protocol requires the guarantee to be met using adequate statistical disclosure control measures. It is a National Statistics principle that for a given statistical output the statistical disclosure control (SDC) measures should not be minimised until the boundaries of the confidentiality guarantee are reached, but rather that the SDC measures for National Statistics should be as comprehensive as can be achieved without undermining fitness for purpose.

#### The Code of Practice states:

4

5(a) Protecting Confidentiality. The National Statistician will set standards for protecting confidentiality, including a guarantee that no statistics will be produced that are likely to identify an individual unless specifically agreed with them.

The Protocol for Data Access and Confidentiality underpins the Code's confidentiality principle, and provides some guidance to the necessary qualification within it. The Protocol sets out the standards for maintaining the Confidentiality Guarantee in the Code. Part 1 of the Protocol states:

Statistical disclosure control methods may modify the data or the design of the statistics, or a combination of both. They will be judged sufficient when the guarantee of confidentiality can be maintained, taking account of information likely to be available to third parties, either from other sources or as previously released National Statistics outputs, against the following standard: It would take a disproportionate amount of time, effort and expertise for an intruder to identify a statistical unit to others, or to reveal information about that unit not already in the public domain.

The terms that add qualification to the 'likely to' phrase in the confidentiality guarantee are: 'intruder', and 'disproportionate time, effort and expertise'.

Other than to distinguish one unit from another for statistical purposes (for example, for data matching or linking exercises for statistical purposes, where

identified data is essential for quality reasons), the statistician or researcher has no interest in the individual statistical unit. In contrast, an intruder is someone who for whatever reason wishes to distinguish one statistical unit in order to treat that unit separately and/or differently from the other statistical units in the dataset, for a non-statistical purpose.

The threats from intruders can vary enormously, and the Protocol can in no way address them all. Therefore consideration of the 'time, effort and expertise' that may be employed by intruders to undo the confidentiality protection of a particular statistic is a responsibility of the authority that publishes the statistic.

Other key terms used in the Protocol are 'identify', and 'to reveal information...not already in the public domain'. The term 'identify' is used frequently in legislation. For example, to distinguish personal census information from census information, the Census Act (1920 as amended) states:

'Personal census information' means any census information which relates to an identifiable person or household.

The term 'identify' can be said to be reserved in legislation for the action of recognising or selecting by analysis the characteristics of a particular person or thing.

The Census Act makes it an offence to disclose any information that relates to an identifiable person or household to another person, without lawful authority. Note the term 'disclose' is reserved in legislation for the action of transferring information (identifiable or otherwise) from one party to another.

Thus the intention of the phrase in the Protocol '...identify a statistical unit to others, or to reveal information about that unit not already in the public domain' is to require that Official Statistics do not allow for the likelihood of an intruder being able to select or recognise by analysis the identity of a particular individual statistical unit, such that the characteristics of the unit could be disclosed to others with confidence as to the correctness of the selection and recognition. The statistic must allow for other information being available to the intruder in addition to this statistic.

Identification of an individual will disclose one or more attributes of that individual. The attributes disclosed are at least those of the table containing the information, and usually the fact of presence in the UK at the time of data capture. This latter may not, in most circumstances, be important, but it is worth noting that solicitors and even the courts have sought this information from publishers of statistics to gain proof of presence in the UK for the purposes of serving divorce petitions, residency claims, etc.

The phrase also requires that where an individual statistical unit is not uniquely identified, but where an intruder can confidently assign attributes to a particular unit that are likely to relate only to that unit, the intruder should

not be able to reveal information about that statistical unit from the statistic that is not otherwise in the public domain.

An individual that can recall their circumstances at the time of data collection (whether survey, census, or administrative collection) will be able to work out which cell in any published table contains their personal information. They will have 'identified themselves', but they will have done so because they know what attributes were provided in the data collection, and they know other relevant information about themselves to assist in the selection by analysis of the cell they are contributing to. Where this cell has a population greater than 1, this generally is not a significant matter for the designers of Official Statistics. Where the cell of self-identification has a population of just 1, or through subtraction or deduction using other available information becomes in effect a population of 1, then this becomes a matter which the Protocol requires to be taken into account.

The Data Access and Confidentiality Protocol requires that publishers of National Statistics are required to consider the issue of self-identification, and in particular to consider and, if necessary, take into account before publication:

- The lawfulness and ethics of publishing statistics that allow selfidentification
- The threat to the credibility of the confidentiality guarantee
- That self-identification will not lead, by subtraction or deduction, to the identification of others
- That the self-identification will not cause the individual substantial damage or substantial distress

#### 3 More on information laws

UK Official Statistics are produced in a complex legal framework. Such is the complexity that this guidance must limit itself to those aspects of the framework that impact upon considerations of protecting confidentiality. This framework can be discussed in three categories:

- · sources of confidentiality rights and protections for individuals
- obligations to prohibit the disclosure of information about individuals, and the authority to supply or acquire individual information to produce statistics
- rights of access to information used to produce statistics

Sources of confidentiality rights and protections for individuals

#### The Human Rights Act

Producers of statistics should be aware that they are obliged in law to establish or maintain the right to a private and family life enjoyed by those subject to the Human Rights Act. Article 8 is interpreted to mean that there is a presumption of no interference by the state in the affairs of the individual and family. Article 8 is not an absolute right, and this privacy may be interfered with - but only to the extent that the interference can be shown to be necessary and proportionate for certain purposes. Where a lawful authority is being relied upon for the interference, the authority must be visible to the individual, used in a manner proportionate to the reason for the interference, and not arbitrary in its effect.

Interference with a private and family life can occur without there being a direct relationship with the individual and/or family. The re-use of administrative information held by a department about an individual must be compliant with Article 8 rights. Producers of statistics that will demonstrably further the proper functioning of the nation's health services should be confident that their interference with Article 8 rights is necessary. Re-using existing information, rather than collecting the information in a new inquiry, is itself an act of proportionality. However, to further minimise the interference, only the information needed for the statistic should be acquired and processed, and the published outputs should guarantee the confidentiality of the information. Where a lawful authority to collect information about individuals for statistics is used, the source of the authority should be readily available, and only the information needed for the stated purposes should be obtained.

#### **The Data Protection Act**

The Data Protection Act recognises the economic imperative that personal data shall be processed. The presumption that personal data shall be processed gives the Act a regulatory function — when personal data are processed, then they shall be processed according to certain principles. An

individual whose personal data are to be processed can expect their information to be processed within the meaning of those principles.

Therefore, an individual whose personal information is to be processed for the production of a health statistic can expect:

- to be informed as to who is determining the manner and purpose of the processing (the 'data controller')
- to be informed as to who if anyone is performing the processing on behalf of the data controller
- to be informed as to the specific purposes of the processing
- that their personal information will be processed only when certain conditions listed in the Act are met
- that the processing of their information does not breach a duty of confidence owed to them, or any other legal prohibition, or limitation on the department's lawful authority to produce the statistic
- to be able to serve notice to cease processing of their information where this
  is causing substantial harm or substantial distress
- to be able to ask for access to their personal information held by the data controller
- to be able to object to any use of their information for the purposes of direct marketing

Exemptions in the Data Protection Act allow that data controllers of personal data used only for statistics do not have to comply with a request for access to information. When processed for statistics, personal data can be kept indefinitely, provided it is being kept for one of the purposes specified to individuals. There is some relief in the Act for the provision of information about the purposes of the processing and the identity of the data controller, particularly where the personal information was obtained from a third party rather than directly from the individual. If providing this information requires a disproportionate effort, then the data controller enjoys relief from this requirement. 'Disproportionate' is not defined in the Act, but is considered a feature of the nature of the data, and the length of time and the cost involved to the data controller in providing the information.

#### Common Law duty of confidence

When a member of the public, or a business or other organisation, provides information to a public authority, a duty of confidence may exist. There is no necessity for an agreement or statement to that effect; whether the duty is owed is a consequence of these circumstances:

 Does the information in question have the necessary quality of confidence? Information not in the public domain, and of some value, will have the necessary quality of confidence. 2. Was the information communicated in circumstances giving rise to an obligation of confidence? The obligation may be express, or implied. It may be reasonable for a person to assume their information as provided to a public authority is communicated in circumstances giving rise to a duty of confidence, whether this was asserted by the authority at the time or otherwise.

The unauthorised use of information that meets these criteria may give cause for an action for damages through a breach of a legal duty. For the action to be successful, it is not always necessary to prove detriment, damage or dishonesty. An unauthorised disclosure may be a disclosure to another public authority or other institution - disclosure into the public domain is not the only unauthorised use that may give cause for an action for damages.

The statistical process transforms information subject to the Human Rights Act, the Data Protection Act and the Common Law duty of confidence. The act of transformation needs to be compliant with these rights and protections, but a statistic that meets the standards of the confidentiality guarantee is information transformed to the extent that it no longer relates to an identifiable individual. The dissemination of such a statistic should be considered free of the obligations to individuals in these sources of law.

Obligations to prohibit the disclosure of information about individuals, and the authority to supply or acquire individual information for statistics

Some data collections add to individuals' general rights and protections by the addition, usually by statute, of a prohibition on disclosure. Statute may absolutely prohibit disclosure, or may limit the disclosure of information to certain recipients, for certain purposes, or a combination of both. The recipients and purposes may be broadly or narrowly defined. These limited conditions for disclosure, together with any express statutory obligations to exchange data with others, can be taken together and described as 'gateways'. For example, the Abortions Regulations contain a prohibition on disclosure of identifying abortions information by the Chief Medical Officer other than to certain prescribed persons for specified purposes. Statisticians are obliged to observe prohibitions on disclosure, but should utilise gateways for exchanging identifying data with other statisticians to produce the best possible statistics.

The production of statistics can be a solution to disseminating information subject to a statutory prohibition on disclosure. Provided the transformation from individual data to a statistic is sufficient, the statutory prohibition may become irrelevant. For example, the dissemination of statistical information derived from abortions notifications allows abortions information to be disseminated to those not able to have the identifying information covered by the regulations disclosed to them. Generally speaking, where the transformation is such that the confidentiality guarantee in the National Statistics Code of Practice is met, then the information should no longer be considered individual or personal information, and therefore should be

considered relieved of obligations to the non-disclosure of identifying information. Authorities should ensure they have the necessary administrative powers (*vires*) to transform identifying data into a statistic. These powers may be express (such as those given to the Registrar General in the Census Act 1920, S5) or, and more usually, will be implied in the statutory or common law powers of the authority. Implied powers are no less valid than express statutory powers.

Where the identifying data originate in another authority, there must be powers to supply and powers to receive such data in the respective authorities. The exchange of identifying data for the purposes of statistics is processing conducted in advance of transformation, and is subject to the information law discussed in this section, and any prohibitions on disclosure. Data sharing for the production of statistics can therefore be complex in law, and more guidance can be found in the ONS publication *Data Sharing for a Statistical Purpose: A Practitioners' Guide to the Legal Framework*.

The UN Fundamental Principles effectively oblige legislation specific to statistics to contain a statutory prohibition on the disclosure of identifying information. The prohibition need not be absolute, but should limit any permitted disclosures to defined statistical purposes only. European law has built this feature into the legal framework for Community statistics:

Article 15. Confidential data obtained exclusively for the production of Community statistics shall be used by national authorities and by the Community authority exclusively for statistical purposes unless the respondents have unambiguously given their consent to the use for any other purposes. (European Union 1997)

The UK does not have a single source of statistical law, but in its absence the relevant law should be interpreted in a manner consistent with this so-called 'one way valve' principle. Data that enter the statistical system, either from statistical inquiries or through the sharing of administrative data, should be used for statistical purposes only, and disseminated only after transformation into a statistic that meets the confidentiality guarantee. The Code of Practice contains this principle, and is an obligation on producers of Official Statistics to ensure the non-disclosure of identifying information for non-statistical purposes.

#### Rights of access to information used to produce statistics

When a request for access to information used to produce a statistic is received, it is important to establish the correct regime for handling the matter. There are three broad types:

- Requests made by an individual for access to information relating only to themselves
- 2. Requests made for access to information relating to identifiable individuals other than themselves
- 3. Requests for access to information that does not relate to identifiable individuals

10

## Requests made by an individual for access to information relating only to themselves

Individuals have a right to ask for access to their own personal information. The Access to Health Records Act 1990 has been heavily amended by the Data Protection Act, and the latter is now the legislative regime for individuals to gain access to their own personal health information. The rights of a Data Subject in the Data Protection Act are limited to their own information only - a request for access to any other person's information becomes a Freedom of Information request. Where the personal health information is being processed only for the purposes of history, statistics and research there is an exemption from the obligation to provide a person with access to their own information. Producers of health statistics should make full use of the exemption, but should always explain why it is being used:

- a statistical dataset is not a definitive source of patient information
- it is only used for statistical purposes
- the individual is of no interest and no measures or decisions relating to individuals will be made with the information, and
- its use can not cause substantial damage or substantial distress to the individual

Officials should appreciate that on rare occasions the decision not to provide a person with access to their individual statistical health information will result in them issuing a Notice to cease processing, on the grounds that the processing is causing substantial harm or substantial distress. Provided all the criteria for what constitutes statistical or research processing are being met, and that there are no non-statistical purposes for the data, the Notice should not be complied with, and the matter should be allowed to go through the appeal process.

# Requests made for access to information relating to identifiable individuals other than themselves

Individuals have a general right of access to information held by public authorities, through the Freedom of Information Act. The Act is now the regime for handling requests for access to identifying information about others. The reasons for not supplying identifying data must be found in the exemptions in the Act. Identifying information used in statistical processes may be exempt information if its disclosure to an applicant would, or would be likely, to constitute one or more of:

- an actionable breach of a duty of confidence
- a breach of a data protection principle
- a breach of a statutory prohibition on disclosure
- prejudice to the effective conduct of public affairs

Officials should ensure that all relevant exemptions are cited when withholding confidential statistical information. While it is good practice to explain a general policy for the withholding of information which is likely to breach the confidentiality guarantee, this must be done in addition to, and not in place of, the language of exemptions from the Freedom of Information Act. The use of exemptions must be determined on a case by case basis, but in all normal circumstances information that does not meet the confidentiality guarantee should be exempt information and withheld. Statisticians should provide relevant qualitative and quantitative justification for the use of an exemption to withhold identifying data.

Where it appears that an exemption does not apply, perhaps because of a police request or court order, the National Statistics Protocol for Confidentiality and Data Access requires that any release of identifying statistical information used for the production of National Statistics should be approved by the National Statistician.

On receipt of a request for information, it is important to consider those cells in published tables disguised/modified for reasons of statistical confidentiality as identifying information in the first instance. The Data Protection Act and the Freedom of Information Act do not make distinctions according to the structure of the information. The reasons for disguising small cells in a publication are ones wholly of statistical policy. However, the reasons for withholding those cells from an applicant under Freedom of Information are those found in the Freedom of Information Act, and should be explained in terms of the Act's exemptions, backed up with reasoning derived from the formulation of the statistical disclosure control policy.

# Requests for access to information that does not relate to identifiable individuals

There are many reasons other than confidentiality protection for disguising cells in published statistical tables, or the non-publication of statistical information. The Freedom of Information Act does not require change to statistical design and dissemination policy. However, where a request for unpublished information is received, the obligation is to provide the information if it is held. Where individuals can not be identified, the relevant exemptions will be few, or none. Producers of statistics should consider the scheduling of publications to pre-empt Freedom of Information requests information scheduled for publication is exempt. There are no exemptions for information of poor quality or of an embarrassing nature, nor even when produced under a faulty methodology. Where information, if released, would, or would be likely to prejudice the effective conduct of public affairs, then that exemption should be considered. Statistical information used for the formulation and development of government policy is exempt, until the decision on government policy has been made. ONS (2004a) has published guidance to the Government Statistical Service on the Freedom of Information Act and statistics which covers these issues in more depth.

#### References

Department of Health (2005) Research Governance Framework www.dh.gov.uk/PolicyAndGuidance/ResearchAndDevelopment/ResearchAndDevelopmentAZ/ResearchGovernance/fs/en

European Union (1997) Council Regulation (EC) No 322/97, EU 'Statistical Law' forum.europa.eu.int/irc/dsis/bmethods/info/data/new/legislation/stat\_law.html

ONS (2004a) Implementing Freedom of Information in Government Statistics and Social Research www.knowledgenetwork.gsi.gov.uk/statnet/statnet.nsf/ RefDocs/SCWE-5U3LSB?OpenDocument

ONS (2004b) National Statistics Code of Practice www.statistics.gov.uk/about/national\_statistics/cop/about.asp

ONS (2004c) Protocol on Data Access and Confidentiality www.statistics.gov.uk/about/national\_statistics/cop/protocols\_published.asp

ONS (2005) Data Sharing for a Statistical Purpose: A Practitioners' Guide to the Legal Framework www.statistics.gov.uk/StatBase/Product.asp?vlnk= 14201&Pos=&ColRank=1&Rank=272

UNECE (1992) E/1992/32/C(47) Fundamental Principles of Official Statistics in the Region of the Economic Commission for Europe www.unece.org/stats/documents/e/1992/32.e.pdf