Access to the National Cancer Data Repository

Version 1.1

Document history

Version No	Date	Author	Notes
1.0	24 May 2010	Michael Chapman	Approved by NCIN Steering Group
1.1	21 Feb 2011	Michael Chapman	Update to NWCIS address

Table of contents

1. Introduction

Overview of the National Cancer Data Repository

1.1 The National Cancer Data Repository (NCDR) comprises a merged dataset of English cancer registration data, linked to an extract of Hospital Episode Statistics (HES). The repository contains over 8.5 million cancer registry records linked to 34 million hospital records, enabling novel and innovative new analyses to be undertaken. The linkage processes have utilised class-leading methodologies, complete with full audits of non-linked records, and detailed quality assurance analyses of linkage performance.

1.2 In addition to the HES linkage, the NCDR has been linked with the General Practice Research Database (GPRD) and approval to link to clinical audit data from the National Clinical Audit Support Programme (NCASP) has been obtained. This document will be revised and updated as required when these and other datasets become generally available.

Legal support for the NCDR

1.3 The current NCDR contains information on England only and is held by the English cancer registries. The information below, therefore, describes the relevant English law. If the NCDR is extended to include data from the other UK nations this document will be updated accordingly.

1.4 In England, Section 251 of the NHS Act 2006 provides the Secretary of State with powers to set aside the common law duty of confidentiality and process patient information for medical purposes without consent. These powers can only be used where it is not practical to gain consent and anonymised information cannot be used. Section 251 replaced Section 60 of the Health and Social Care Act 2001, which provided the same powers.

1.5 A regulation made under Section 60[1] (and continuing to have effect under Section 251) permits cancer registries to receive patient identifiable data on those referred for the diagnosis or treatment of cancer without the need for informed consent and it permits registries to process this data for the medical purposes stipulated in the regulation.

Confidentiality of the NCDR

1.6 Cancer registration data are released subject to the confidentiality and disclosure guidelines, as agreed by the UK Association of Cancer Registries (UKACR). This document is based on these policies[2,3], the UKACR Executive's discussions on the release of national cancer registration data and the English cancer registries' agreement with the NHS Information Centre on the release of HES data.

[1] Regulation 2 of the Statutory Instrument (SI) on confidentiality – No. 1438, The Health Service (Control of Patient Information) Regulations 2002

[2] UKACR (2003). Policy on release of identifiable patient information. Retrieved 14th July 2009 from http://82.110.76.19/confidentiality/policies.asp

[3] UKACR (2003).Policy on release of potentially identifiable patient information. Retrieved 14th July 2009 from http://82.110.76.19/confidentiality/policies.asp

2. Requirements for access

2.1 The criteria for accessing NCDR data will vary based on whether the request is for identifiable data and whether information from one of the linked datasets is required. The UKACR policies on release of identifiable patient information and release of potentially identifiable patient information contain information to help determine the nature of the request. The eligibility requirements for each type of data are detailed below.

Anonymised data

2.2 Fully anonymised data may be requested from cancer registries without requiring other approvals (but see the relevant UKACR policies to confirm that the data are not potentially identifiable). Before applying for such data applicants should check the summary statistics and other information available from cancer registries, the cancer e-Atlas[4] and the National Cancer Intelligence Network (NCIN) to ensure that the data are not already publically available. NHSnet users should also check that the required information is not available through the National Cancer Information Service (NCIS)[5].

Identifiable data

2.1 When dealing with requests for patient identifiable data, registries must assess each request carefully and on merit, to ascertain whether or not patient identifiable data are really necessary. If not, anonymised data must be supplied, following the relevant UKACR guidelines.

2.2 In general, identifiable data can be released to:

- Authorised personnel in approved organisations; *or*
- Those holding either (A) Patient Consent or (B) approval from the Ethics and Confidentiality Committee (ECC) of the National information Governance Board for Health and Social Care (NIGB). Research studies will also require approval from a Research Ethics Committee (REC).

Authorised personnel in approved organisations

2.3 The UKACR policy on 'Disclosure of identifiable data by cancer registries' provides full details of approved organisations and which personnel are authorised to receive identifiable data from the English cancer registries.

Other applicants

2.4 All other applicants for identifiable data require either evidence of informed consent from all individuals whose information is requested or support from the ECC under Section 251 of the Health and Social Care Act (2006).

2.5 Research projects will also require approval from the appropriate REC[6].

[4] See http://www.ncin.org.uk/eatlas/

[5] See http://www.ncin.org.uk/ncis/

[6] NRES provides guidance on the difference between Research, Clinical Audit and Service Evaluation (see NRES (2008). Defining research. Retrieved 14 July 2009 from www.nres.npsa.nhs.uk/applications/apply/is-your-project-research/)

Potentially identifiable data

2.6 The disclosure of anonymised but individual records or tabular data to a low level of aggregation (including some low cell counts) poses a very small but theoretical risk of identifiability when combined with other existing knowledge of the data recipient(s), or knowledge obtained from a different source.

2.7 The issues are complex and the range of scenarios (in terms of combinations of variables, levels of aggregation, and availability to the data recipient of additional knowledge) is almost unlimited: it is not possible to be unduly prescriptive without being unjustifiably restrictive. The UKACR policy on release of potentially identifiable patient information provides a series of pragmatic guidelines designed to guide registries in coming to decisions about the release or publication of potentially identifiable data. All releases will be made in line with this policy.

2.8 Requests for potentially identifiable data will be scrutinised and approved (or not) by the director of the relevant cancer registry. They should consult the relevant registry advisory group as appropriate and, if necessary, the Caldicott Guardian of their host organisation. The applicant may also be referred to the ECC and/or a REC.

HES linked data

2.9 While cancer registries can provide analyses based on HES linked data, most requests for linked datasets for analysis by third parties must currently be approved by the NHS Information Centre. Requests for certain sensitive HES data items may also require additional approvals as described in the HES protocol[7].

2.10 Further guidance on when approval is necessary has been issued to cancer registries[8]. The cancer registry dealing with the request will advise applicants on the need for approval and handle this with the Information Centre.

GPRD linked data

2.11 GPRD is committed to ensuring that GP and patient confidentiality is maintained at all times. To protect this patient and practice anonymity, no identifiable information can be released for linkage to GPRD, even if consent or other permissions are held.

National clinical audit (NCASP) linked data

2.12 [Any limitations on the release of NCASP linked data will be added here]

Application for data from multiple sources

2.13 Currently data may be supplied from only two linked sources at a time (e.g. registry-HES or registry-GPRD but not registry-HES-GPRD). Any requests for linked information from more than two sources are likely require specific approvals from the NIGB's Database Monitoring Sub-Group (DMSG) or the ECC.

[7] NHS Information Centre for health and social care (2009). HES Protocol. Retrieved 14th July 2009 from http://www.hesonline.nhs.uk/

[8] Thames Cancer Registry (2007). Thames Cancer Registry guidance on the release of HES data.

3. Applying for access

Where to apply

3.1 For requests covering a single tumour type, application should be made to the lead cancer registry for that tumour type. Requests covering multiple tumour types should initially be directed to the applicant's local cancer registry, although the application may be passed to another cancer registry based on expertise and workload (see 4.2 below).

3.2 Applicants who wish to access regional cancer registration data should apply to their regional cancer registry under normal UKACR procedures.

3.3 Contact details for the English cancer registries and a list of their lead areas are given in Appendix A. Appendix B summarises where applications should be directed and which group will handle them.

3.4 Applicants are strongly encouraged to discuss their requirements with the cancer registry in advance. Registry staff will be able to provide advice on the suitability of the NCDR for the proposed use.

GPRD linked data

3.5 Applicants for GPRD linked data should apply to GPRD[9] in the first instance, stating that they wish to use the cancer registry linkage. They are, however, encouraged to discuss their request in advance with the appropriate cancer registry to ensure that the registration data they require is suitable for their purpose.

Information to be supplied by applicants

3.6 Applicants should provide a detailed protocol justifying their data requirements together with a completed data request form and return this to the appropriate cancer registry.

Data items and interpretation

3.7 Applicants should request only the data items required for their work. Although there are many data items included in the NCDR, cancer registries are obliged to restrict data release to the minimum required for each study, at the highest level of anonymisation possible.

3.8 The list of data items is only intended as a starting point in a dialogue with cancer registry staff, who can provide advice on the needs of each study and whether particular data items are complete or robust enough for analysis.

3.9 Applicants should also consult the cancer registry on the interpretation of data – for example, changes in coding practice over time can have large impacts on data, which may be artefactual and not epidemiological.

Timing for requests for access

3.10 Requests for information can be made at any time. Requests for information will be responded to within 20 working days of receiving a completed application. Requests which require additional information processing or analysis may take longer. Applicants can help ensure that their request is dealt with promptly by supplying all relevant

[9] See http://www.gprd.com/isac/

information at the time of their application and discussing their requirements with cancer registry in advance.

Requests made prior to securing funding or required approvals

3.11 Where it is required to secure funding, the NCIN coordinating team can arrange a letter stating that the required data are available and agreeing in principle to supply them (subject to the necessary approvals). If this is required it should be specified in the application.

4. Processing applications

4.1 On receipt of an application, registry staff will check that all necessary information has been supplied and that the appropriate approvals for the work are either already in place or are being sought. If requests are incomplete the applicant will be contacted and asked to resubmit the application with the required information.

4.2 Applications for information on a single tumour site will be handled by the lead registry for that area. Applications for information covering multiple tumour sites will be passed to the UKACR Analysis Subgroup, requesting that one of the English cancer registries agrees to handle it. It is expected that the group will be able to assign most requests to a registry with particular expertise or interest in the area covered by the request; where this is not possible, requests will be assigned based on a rota.

4.3 Input into extracts requiring specialist knowledge of the NCDR may be provided by the Northern and Yorkshire Cancer Registry and Information Service (NYCRIS) or Thames Cancer Registry (TCR). Requests for GPRD linked data will be handled initially at NYCRIS; this arrangement will be reviewed once the processes for releasing GPRD linked data are fully established. Appendix B summarises how requests will be allocated to registries.

Approval by the cancer registry director

4.4 For national data that are identifiable or potentially identifiable, it is the responsibility of the director of the cancer registry processing the request to ensure that any disclosure is in line with the approved UKACR policies and to authorise the disclosure if they consider it necessary.

4.5 The registry director should ensure that the other English cancer registries are consulted:

i. to ensure that the records of any patients who have requested the deletion of their data have been removed from the national repository or appropriately flagged; *and*
ii. identify any other issues (e.g. data quality or particular risks of identification) that may influence the decision to release the data.

4.6 Depending on the nature of the request, the director of the registry processing the request may require a positive response from the other English cancer registries or may consider that no response within a reasonable time (not less than twenty working days) indicates agreement.

4.7 Currently this consultation is expected to take place through an email to the directors and analysis leads of the other English cancer registries. The NCIN coordinating team will investigate other methods for ensuring that applications are tracked and that registry directors are aware of all releases.

Approval by the NHS Information Centre (HES linked data only)

4.8 The guidance issued by TCR to the English Cancer Registries sets out when releases of HES linked data require approval from the NHS Information Centre.

4.9 Where required, this will be handled through the cancer registry processing the request. They should discuss the request and seek approval from Louise Dunn on information governance issues and Dean White on questions relating to re-use of data.

Approval for GPRD linked applications

4.10 Applications to GPRD will be considered by the Independent Scientific Advisory Committee (ISAC) of the Medicines and Healthcare products Regulatory Agency (MHRA). In parallel with this process, a subgroup of the NCIN Steering Group will arrange for a light touch review of the application to ensure that:

i. the data requested is suitable to answer the question posed;
ii. the question is an important one to address;
iii. the release is compliant with UKACR policy; *and*
iv. the applicants have made appropriate contacts within the registry community (e.g. with the lead registry if the proposal relates to a particular cancer site).

4.11 If both groups approve the request it will be passed to NYCRIS for processing.

Informing Caldicott Guardians

4.12 Where required, details of releases will be provided to the cancer registry's Caldicott Guardian, who may consult with or inform counterparts in other organisations.

Screening for technical / scientific / clinical merit

4.13 Applications will initially be screened to ensure their technical, scientific and clinical merit (as appropriate) by the cancer registry handling the request. Where appropriate this will be referred to the NCIN Scientific Advisory Group or the relevant NCIN Site Specific Clinical Reference Group.

Prioritisation of requests

4.14 Where necessary requests will be prioritised on the basis of expected benefit to NHS users and the public.

Appeals

4.15 Applicants may appeal a discussion not to release information requested from the NCDR to the UKACR Executive. Any appeals should be made in writing to the cancer registry dealing with the application within 28 days and will be considered at the next meeting of the Executive.

4.16 Where the MHR ISAC advises against the release of GPRD data, applicants may appeal in writing (addressed to the ISAC Secretary) within 28 days[10].

[10] See http://www.gprd.com/isac/

5. General conditions of access to the NCDR

5.1 All recipients of identifiable or potentially identifiable data will be required to sign a confidentiality form as part of their application.

5.2 The following conditions of access apply to all releases of data from the NCDR unless explicitly stated otherwise. Specific conditions apply to HES, GPRD and national clinical audit linked data; these are highlighted in section 6 below.

Limitation of use and onward transfer

5.3 The intended use(s) of the data should be stated clearly and justified in the application. Recipients will be required to agree that identifiable or potentially identifiable data will not be used for any other purpose and that no attempt will be made to link these data to other data sets unless this is agreed in advance with the cancer registry.

5.4 Recipients of identifiable or potentially identifiable data must not pass this to third parties or release it into the public domain.

Data identifiability

5.5 Recipients must agree that no attempt will be made to identify information pertaining to particular individuals or to contact individuals (unless patient consent has been obtained via the patient's clinician).

5.6 Any results which are disclosed must not identify any individual and any results which are disclosed into the public domain[11] must not show any potentially identifiable information.

Publication and acknowledgement

5.7 Any public domain[11] reports or papers resulting from analyses of identifiable or potentially identifiable data should be shared prior to publication with the cancer registry supplying the information. The cancer registry will provide any requests for changes to be made prior to publication within twenty working days of receipt.

5.8 Where a request may lead to release into the public domain of information relating to variations in outcomes or patient management, the cancer registry should discuss plans for publication with the requestor at the proposal stage. Advice on this may be requested from NCIN's Site Specific Clinical Reference Groups at the discretion of the director of the registry handling the request.

5.9 All publications using information from the NCDR should acknowledge the source of the information as the ["National Cancer Data Repository (NCDR) provided by the English cancer registries on behalf of the National Cancer Intelligence Network"]. This should be supplemented as required with acknowledgement of other linked data sources (see section 6 below).

5.10 Where cancer registry staff have made substantial contributions to the work, their names should be included as authors in the publications or presentations and they should be given full authorship rights.

[11] Publication of data on a website and in unrestricted circulations of reports or documents containing data should be regarded as being in the public domain.

Opt out of patients from the cancer registration scheme

5.11 Under UKACR policies[12], three options are offered to cancer patients who wish to opt out of the cancer registration scheme:

> **Option 1:** the registry will retain the record but flag it to ensure identifiable data are not released to any research projects involving patient contact
>
> **Option 2:** the registry will delete all clinical details currently held but will keep a record of the patient's ID in order to delete any information received at a later date
>
> **Option 3:** the registry will delete all records currently held, including the patient's ID - this will mean that any data received in the future will not be able to be linked to the request for deletion and so will be processed in the normal way.

5.12 Where a patient selects options 2 or 3, the cancer registry dealing with *the patient's request* should attempt to identify any recipients of identifiable data about that patient from the NCDR. Recipients should be contacted and instructed to delete the record from their data and destroy any paper records. In the case of option 1, recipients of identifiable data about the patient should be instructed to flag the record on their live database as not to be used for research for patient contact.

Data security and destruction

5.13 Identifiable or potentially identifiable data should be kept securely for the period of time that can be justified by the stated purpose, and then destroyed. Once the data have been destroyed, this should be confirmed to the cancer registry that supplied the data.

5.14 Advice on appropriate electronic and physical security measures may be sought from the cancer registry.

Compliance with the conditions of access

5.15 It is the recipient's responsibility to ensure that the necessary information governance and security procedures are in place and that all information supplied from the NCDR is used in accordance with these conditions.

5.16 Recipients found to be in breach of the agreement under which data has been supplied will be denied future access to the NCDR and their institutions and funders informed. Breaches of confidentiality may also expose recipients and their institution to legal penalties.

Transparency

5.17 A database of all releases from the NCDS will be maintained and a summary of this information (including a summary of the information released, the reason for the release and the recipient of the data) will be published on the NCIN website.

5.18 Recipients are requested to provide details (and where appropriate a copy) of any publications or reports based on the data for inclusion in this database.

[12] UKACR (2007). Policy on responding to requests from patients to delete their identifiable information. Retrieved 14th July 2009 from http://82.110.76.19/confidentiality/policies.asp

Fees

5.19 Requests from NHS and academic organisations that require only simple processing will generally be met without charge.

5.20 Requests from other organisations and those that are complex or time consuming may incur a fee to cover the cancer registry's costs of retrieving and processing data. The cancer registry handling a request will inform applicants of whether a fee will be charged and the total cost.

5.21 Additional fees may be incurred for the processing by third parties required to provide some linked datasets. Details of these are given in section 6 below.

6. Specific conditions of access to linked datasets

Additional conditions for accessing HES linked data

6.1 All publications based on HES Data must comply with the HES Protocol in addition to the conditions specified above.

6.2 Publications making use of HES linked data must acknowledge that "The NCDR contains data from Hospital Episode Statistics (The NHS Information Centre for health and social care)". They must also cite the NHS Information Centre's copyright as follows:

"Copyright ©[YEAR], Re-used with the permission of The Health and Social Care Information Centre. All rights reserved."

6.3 Depending on the type of HES information supplied other footnotes and caveats may be required. Details of these will be supplied by the cancer registry.

Additional conditions for accessing GPRD linked data

6.4 In addition to the conditions above, any use of information released by GRPD must comply with the licensing conditions agreed with GPRD.

Additional conditions for accessing national clinical audit linked data

6.1 [Any additional conditions associated with NCASP linked data will be inserted here]

Appendix A: Contact details and lead tumour sites for the English Cancer Registries

Contact Details	Lead Tumour Areas
Eastern	
Eastern Cancer Registration and Information Centre Unit C - Magog Court, Shelford Bottom, CAMBRIDGE, CB2 4AD www.ecric.org.uk/	Brain / CNS
North West	
North West Cancer Intelligence Service The Palatine Centre 63-65 Palatine Road Manchester, M20 3LJ www.nwcis.nhs.uk/	Children & TYA (with Childhood Cancer Research Group)
Northern and Yorkshire	
Northern and Yorkshire Cancer Registry and Information Service Level 6 St James's Institute of Oncology Bexley Wing St James's University Hospital Beckett Street LEEDS, LS9 7TF www.nycris.org.uk/	Colorectal Haematological
Oxford	
Oxford Cancer Intelligence Unit 4150 Chancellor Court, Oxford Business Park South OXFORD, OX4 2GX www.ociu.nhs.uk/	Head and neck
South West	
South West Public Health Observatory South West Cancer Intelligence Service Grosvenor House 149 Whiteladies Road BRISTOL, BS8 2RA www.swpho.nhs.uk/	Skin Urology
Thames	
Thames Cancer Registry King's College London School of Medicine 1st Floor, Capital House 42 Weston Street LONDON, SE1 3QD www.tcr.org.uk/	Lung Upper GI (inc HPB)
Trent	
Trent Cancer Registry 5 Old Fulwood Road	Gynaecological

SHEFFIELD, S10 3TG www.trentcancer.nhs.uk/	
West Midlands	
West Midlands Cancer Intelligence Unit Public Health Building The University of Birmingham Edgbaston BIRMINGHAM, B15 2TT www.wmpho.org.uk/wmciu/	Breast Sarcoma
Childhood Cancer Research Group	
Childhood Cancer Research Group 57 Woodstock Road OXFORD, OX2 6HJ www.ccrg.ox.ac.uk/	Children & TYA (with North West Cancer Intelligence Service)

Appendix B: Where should applications be directed and who will deal with them?

Data required		Apply to:	Request processed by:
Registration data	Regional	Local cancer registry	Local cancer registry
	National - single tumour site	Lead registry for tumour site	Lead registry for tumour site
	National - multiple tumour sites	Local cancer registry	Assigned to appropriate registry
HES linked registration data	Regional	Local cancer registry	Local cancer registry (TCR or NYCRIS for extracts requiring specialist knowledge of the NCDR)
	National - single tumour site	Lead registry for tumour site	Lead registry for tumour site (TCR or NYCRIS for extracts requiring specialist knowledge of the NCDR)
	National - multiple tumour sites	Local cancer registry	Assigned to appropriate registry (TCR or NYCRIS for extracts requiring specialist knowledge of the NCDR)
GPRD linked registration data		GPRD	NYCRIS

Notes:

'Regional' refers to an area substantially covered by a single cancer registry. 'National' includes any requests requiring data from more than one cancer registry.