Data Protection Impact

Assessment

National COVID-19 Chest Image Database

Template

Version

number: 1

Published

Date:

Prepared by: Corporate

Information Governance

Version Control			
Reference	Change	Date	Approved
1.0	Initial Draft		
1.1	Updates for IG		
1.2	Updated with Corp IG Comment	7 Apr 2020	
1.3	Final for Approval	9 Apr 2020	
1.3.1	Extended age range of data collection	12 May 2020	
2.0	Linkages to other databases Inclusion of patients without imaging in the acute setting	1 Sept 2020	

Version Control

Data Protection Impact Assessment
National COVID-19 Chest Imaging Database

Administrative information

NHS England	
Your name	Dominic Cushnan
Your team and directorate	Innovation, NHSX, NHSE
Your location	Quarry House
Your telephone number	07590 288865
Your email address	dominic.cushnan@nhsx.nhs.uk

NHS Improvement	
Your name	Dominic Cushnan
Your team and directorate	Innovation, NHSX, NHSE
Your location	Quarry House
Your telephone number	07590 288865
Your email address	dominic.cushnan@nhsx.nhs.uk

Purposes

Fully describe what is the purpose of the project and how is the processing of information necessary to that work?	NHSX would like to create a national database of chest X-ray, CT and MR images and other relevant patient information that enables the development and validation of automated analysis technologies that may prove effective in supporting COVID-19 care pathways, and that accelerates research projects to better understand the disease. In addition to the above uses, the database will be utilised for teaching purposes to support the training of radiologists and experts. The processing of de-identified patient data (health records, X-ray, CT and MR images) are necessary for the creation of the database. The use of this data by researchers and technology developers will be necessary for the development of support tools in response to the COVID-19 crisis. Data may be annotated and augmented by experts after being collected using a web portal. This user interface would display images and data to the expert, and request an annotation or action. Responses provided by the experts would

	subsequently be stored in the database alongside the images and data. Annotated data may be used in creating different types of software tools (including AI ones) that process augmented information. Annotated data may also be used to simulate the interaction between software and users in the validation of products (including AI ones) such as clinical decision-support systems. Data may be utilised by external auditors for the validation of efficacy and safety of AI tools for COVID-19 chest imaging.

Nature of the data

Will the processing involve anonymised information[1]?	No
Will the processing involve pseudonymised personal data?	Yes
Will the processing involve fully identifiable personal data?	No

Assets

Does the proposal involve creating a new information asset?	Yes, the proposal involves the creation of a new information asset, in the form of a national database. Imaging data (X-Rays, CTs and MRs) will be a copy of the data normally acquired by participating sites. A tailored sheet of additional patient information will also be created for the purposes of this project.
Does the proposal involve processing data held on an existing information asset or assets?	Yes. This will involve processing X-Ray, CT and MR images already held in the systems of each participating site.

`

[1] anonymous information is information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable

Is/are the asset owner(s) aware of the proposal	Yes.

What is the timeframe for the project/programme/initiative? (Please include commencement dates and any foreseen end dates)

The project will commence immediately upon approval of the DPIA.

No end date has been defined yet. Data will be made accessible to authorised and permitted researchers to undertake specific research projects to address needs in the detection and stratification of COVID-19 that will benefit from combined clinical and imaging data. Any access to the data and licences to use developed to support the COVID-19 pandemic, and will cease when the COVID-19 COPI (COVID-19 – Notice under Regulation 3(4) of the Health Service Control of Patient Information Regulations 2002) ceases effect

Controllers[2]

NHS England	Joint with DH, and also joint with NHS Scotland for Scottish data only
TDA	No
Monitor	No
NHS Digital	No
Other (Please do not include any third party that we are contracting with to process personal data for us as a processor.)	

Screening questions

Does the proposal involve any of the following – drop down list to include: ● NCDR ● Pseudonymised by NHS Digital ● Aggregate data ● Anonymised data	Yes, it involves the collection of pseudonymised data.
Has processing of this nature already been captured and considered within a previous DPIA? If so, link to reference number	No

`

[2] 'controller' means NHSE, alone or jointly with others, determines the purposes and means of the processing of personal data

Will the processing involve a large amount of personal data (including pseudonymised personal data) and affect a large number of data subjects?	Yes, the processing will involve the collection of an extensive amount of patient data (pseudonymised) and will affect a large number of subjects. As the number of patients suspected of having COVID-19 will increase, so will the number of patients whose data (X-Ray, CT or MR images and relevant healthcare records) will be included in the database.
Will the project involve the use of a new technology(ies) which might be perceived as being privacy intrusive? i.e. using biometrics, facial recognition, Artificial Intelligence or tracking (such as tracking an individual's geolocation or behaviour)?	Yes, the project will involve the development of analysis tools and Artificial Intelligence tools that comply with all current regulation. However, it is unlikely that individuals would consider this an intrusion of privacy, as patients' privacy will be safeguarded at every stage and data will only be used for creating tools to support the response to COVID-19.
Will the processing introduce or make use of a new platform not currently in use?	Yes, a new platform will be developed for the purposes of this project, to host data and provide the users with the capabilities to access the data. Furthermore, a cloud-based environment that supports the creation of servers and databases, whilst providing productivity tools for software programming, will be utilised for the validation of AI tools. The platform will be provided by Faculty. (Faculty is a pre-existing contracted supplier, see Supplier Information section)
In the absence of proper controls is there the risk that the processing may give rise to discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of personal data protected by professional secrecy (e.g. health records), unauthorised reversal of pseudonymisation, or any other significant economic or social disadvantage?	No, the personal data is limited to the medical information required for the purposes of this project and would not give rise to any of the scenarios.
Does the proposal introduce difficulties in ensuring that individuals are informed or able to exercise their information rights?	Yes. Under the current circumstances it will not be possible to ensure all patients are informed through posters and

	leaflets. The use of patient information in this instance is covered by The Health Service (Control of Patient Information) Regulations, under the “Communicable disease and other risks to public health” section. Given the tight deadlines, this will allow us to ensure data is collected and provided to researchers and developers as soon as possible. For Scottish data, the use of data is covered by PBPP application 2021-0018 (as COPI notices do not apply in Scotland).

Will there be processing of genetic data, data concerning health, sex life, racial or ethnic origin, biometric data, political opinions, religion or philosophical beliefs, or trade union membership?	Yes, there will be collection and processing of data concerning racial or ethnic origin.
Will there be processing of data concerning criminal convictions and offences or related security measures?	No
Will the project involve the targeting of children or other vulnerable individuals for marketing purposes, profiling or other automated decision making?	No
Will the processing result in you making decisions or taking actions against individuals in ways which can have a significant impact on them? e.g. decisions about an individual's access to a product, service, opportunity or benefit, or recruitment aptitude test based on automated decision making (including profiling)?	If the analyses and development of support tools will be successful and result in utilisable technologies, then the use of those technologies will assist doctors and healthcare workers in managing and diagnosing patients, therefore having an impact on their decision making process and on individual patient's care.
Will there be a systematic monitoring of a publicly accessible area on a large scale (e.g. CCTV)?	No
Will the processing include any data matching e.g. the combining, comparing or linking of personal data obtained from multiple sources?	Yes. The processing will include the linkage of the data in the database to additional clinically relevant information held in other research and clinical databases. This DPIA covers the linkage with the ISARIC database for the collection of the NCCID clinical variables through the ISARIC database, avoiding duplication of activities for research nurses; the linkage to the segmentation dataset is to increase the completeness of the ethnicity information.
Will personal data about individuals be shared with other organisations or people who have not previously had routine access to the data?	No
Will the project/proposal use personal data about individuals for a purpose it is not currently used for or in a new way?	Yes, the project will use routinely collected data for additional purposes, i.e. development of technology tools and advancement of research on COVID.
Will the project require you to contact individuals in ways which they may find intrusive? i.e. telephoning or emailing them without their prior consent.	No

Are you using a Data Processor/third party supplier or is a service/processing activity being transferred to a new supplier/organisation (or re-contracted) at the end of an existing contract?	No

NB. If the answer to any of the above questions is Y, please complete the rest of the form. If all of the screening questions are answered N, the local IG team must still sign off the DPIA.

Where the information will include the processing of personal data, please continue.

Personal data[3]

Why would it not be possible to do without personal data?	Personal data is required for understanding clinical patterns of the disease and to improve the efficacy and accuracy of technology tools
What are the required personal data? Please itemise them or supply a dummy sample, blank forms, screenshots from the prototype system etc.	Age, Gender, NHS/CHI number, ethnicity and the clinical information contained in the sheet in attachment (see COVID-19_NCCID_covid_positive_data_template_v1_6.xlsx).
Please confirm that this is the minimum amount of personal data that is necessary.	Yes, this is the minimum amount of data that is required for the database to be valuable to data users and technology developers.
Would it be possible for NHSE to use pseudonymised personal data for any element of the processing?	Yes, the NHS/CHI number will be pseudonymised.
If Y, please specify the element(s) and describe the pseudonymisation technique(s) that we are proposing to use.	The patient pseudonym will be created by a lossful encoding algorithm and complex salt which will produce an encoded pseudonym which allows the linking of the clinical data with the images. No patient identifiable information will leave any of the participating sites.

`

[3] 'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an

identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person

Scale and constituency(ies)

What is the scale of the processing (i.e. (approximately) how many people will be the subject of the processing)?	Patients of hospitals within the UK. 20,000 - 40,000 individuals
Please describe the constituency(ies).	Patients suspected of having COVID-19 who underwent a COVID-19 swab test. A number of these patients will be selected for their data to be uploaded to the database. Data on children will also be collected. Data of children below the age of 12 will not be shared with external users until there is a sufficient volume of similar patients so that an individual cannot be identified. Chest imaging data (XRay, CT or MR) for these patients may or may not be collected in the acute setting. It is likely of value to store historic imaging in patients who were tested for COVID but who underwent no imaging in the acute setting. We will have clinical data on these patients at the time of assessment. Comparing the historic/pre-COVID imaging in these patients who may have undergone mild COVID-19 infection with the pre-COVID imaging in patients who underwent severe infection (determined by need for acute imaging and/or ITU admission) could allow delineation of imaging-based biomarkers that indicate a) susceptibility to severe COVID-19 infection (in those undergoing acute imaging), b) imaging features protective against severe COVID infection developing. The potential for combining clinical variables with imaging biomarkers in the population where no acute imaging was needed would improve the robustness and generalisability of risk prediction models and reduce the risk of biases from only analysing patients with severe disease.

Outcomes

What will be the effects of the processing (i.e. what actions/decisions will result from the processing)?	NHS E and DH authorised and permitted researchers and technology developers will be able to utilise the data for their research projects and development of tools in response to the COVID-19 crisis. The processing of the data will result in deeper clinical understanding of the disease and in the development of tools that will help in: - Speeding up patient triage processes, supporting instant risk stratification and cohorting of chest presentations - Analysing the appearance/features of chest X-rays, CTs and MRs in order to characterise disease progression, outcomes and complications - Identifying other issues with the patient’s lungs that may have prompted an urgent care attendance, and may complicate the patient triage for COVID-19

Joint working controllership relationship and bases for lawful processing (if sole controller then proceed to next section)

Which controllership scenario(s) below apply(ies)?		**Yes/No**	**Legal basis GDPR Article 6**				
			NHS E	**TDA**	**Monitor**	**NHSD**	**Other**
1. Joint controllers – aligned exercise of specific statutory functions	The Parties have separate statutory functions as the basis for conducting activities and processing, but the functions are related and their exercise is to be aligned. For example, processing to support • NHS England’s functions in respect of the performance assessment of CCGs and giving directions to CCGs • NHS Improvement’s functions in respect of the oversight and regulation of NHS trusts and Foundation Trusts.	es	A6(1)e performance of a task carried out in the public interest				- DH - NHS Scotland (for Scottish data only)
2. Joint controllers – general powers and corporate governance arrangements	The Parties collaboratively process personal data in the exercise of their general powers. For example, processing to support •The appointment of joint executive positions on the board and at a senior level •Shared HR service •Planning for operational management of integrated teams •Line management of integrated teams •The establishment of a shared secretariat for servicing boards and committees.						
3. One Party is a Controller, supported by staff employed by any of the other Parties	One Party alone is responsible for determining the purpose and means of Processing to exercise its own functions and consequently it is the sole Controller. An employee of another of the Parties who assists with the Processing under the guidance, direction or supervision of the sole Controller is acting as an agent of the Party which has the function and which is the Controller.						

4. Data sharing –	Information is shared between						

with each organisation as a separate controller	Parties as individual data controllers. Personal data is disclosed by one Party to another for the latter's discrete purposes. The Parties undertake their own functions, but with the support of either or both of the others in a manner which involves the sharing of personal data						
Disclosing controller	NHS England / TDA / Monitor						
Recipient controller	NHS England / TDA / Monitor						
5. Processor	One or more of the Parties acts as a processor for one or more of the other Parties. The processing Party does not need a legal basis, relying on that of the controller. For example One of the Parties remains as data controller responsible for determining the purpose and means of processing, but the processing of data on itself is undertaken by another Party.						
Controller(s)	NHS England / TDA / Monitor						
Processor	NHS England / TDA / Monitor						

Purpose(s) and legal basis(es) of the processing (complete only if sole controller)

(Please tick all that apply.)	
Is the processing necessary for a task that is within NHSE's remit as a public authority? (please specify below)	Yes
(This is applicable for much of NHS England's processing of personal data using its statutory powers) GDPR Article 6(1)(e) Covid Purpose supported by COPI Reg's Notice to identify Chest X-ray or CT or MR Image for inclusion.	
Is NHSE under a legal obligation to carry out the processing? (please specify below)	No
(e.g. NHSE is obliged to disclose employees' payments and deductions to HMRC on or before each payday) GDPR Article 6(1)(c)	
Is the processing necessary for the arrangement or fulfilment of a contract between NHSE and the subject(s) of the personal data? (please specify below)	No
(e.g. recruitment and management of staff contracts) GDPR Article 6(1)(b)	
Will we be seeking, and recording, freely given, specific and informed consent[4] to the processing? If so, please supply a copy of the draft consent form.	No
(Note that the definition of consent for GDPR purposes is very rigorous. If we are using another legal basis e.g. 6(1)(e) GDPR Article 6(1)(a)	
Is the processing necessary in an emergency situation to protect the life or safety of any person? (please outline below)	No

`

[4] 'consent' of the data subject means any freely given, specific, informed and

unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her – this must be demonstrable by NHSE

(NB This basis should be used only where the processing cannot be based on another legal basis.)	
(e.g. NHSE might be asked to supply non-confidential personal data (e.g. name and address) about patients to third parties in order to facilitate emergency care following a civil or natural disaster) GDPR Article 6(1)(d)	
Is the processing necessary in the legitimate interests of NHSE or a third party?	No
(legitimate interests is not available as a basis for our statutory tasks, but may be used for incidental activities) GDPR Article 6(1)(f)	
In this last case, please specify the legitimate interests and explain why and how they are not in conflict with the interests or rights and freedoms of the subjects of the personal data.	

Special categories of personal data

Will the processing Involve personal data about: (Please tick all that apply.)	
● racial or ethnic origin	Yes
● political opinions	N/A
● religious or philosophical beliefs	N/A
● trade union membership	N/A
● genetic data[5]	N/A
● biometric data[6]	N/A
● data concerning health[7]	Yes

`

[5] 'genetic data' means personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person and which result, in particular, from an analysis of a biological sample from the natural person in question

[6] 'biometric data' means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data

[7] 'data concerning health' means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status

● data concerning the sex life or sexual orientation of the data subjects	N/A

If there are no special categories of data processed, please skip the following section and proceed to the ‘Common law duty of confidentiality’ section…

Legal basis(es) for special category personal data

Legal basis	Personal data to which this legal basis relates: (Please indicate applicable categories from previous section, or “all”.)
● explicit consent	GDPR Article 9(2)(a)
No	
● required in the field of employment, social security or social protection law (please specify below)	GDPR Article 9(2)(b)
No	
● necessary in an emergency situation to protect the life or safety of any person where the data subject cannot consent (please specify below)	GDPR Article 9(2)(c)
No	
● data subject has put the personal data in the public domain	GDPR Article 9(2)(e)
No	
● necessary for legal claims or to the Courts (please specify below)	GDPR Article 9(2)(f)
No	
● necessary for reasons of substantial public interest (please specify below)	GDPR Article 9(2)(g)
No	
● necessary for health or social care purposes (please specify below)	GDPR Article 9(2)(h)

Yes: Covid-19 purpose

● necessary for public health (please specify below)	GDPR Article 9(2)(i)
Yes: Covid-19 purpose	
● necessary for archiving in the public interest, scientific or historical research purposes or statistical purposes (please specify below)	GDPR Article 9(2)(j)
No	

Common law duty of confidentiality

Are any of the data subject to a duty of confidentiality (e.g. clinical records, OH details, payroll information)? If so, please specify them.	Yes, patient clinical records.	
Where it is planned to disclose such data, what are the grounds for doing so?	● consent ● safeguarding ● other overriding public interest - please specify ● legal duty or permissive power e.g. s251 support – please specify (e.g. court order)	Other overriding public interest: the data will be used to accelerate research and development of technological tools to support the COVID-response. Permissive powers: COPI Notice CLDC, GDPR/DPA(18) and Public interest/Public Health legal cover. For Scottish data only, as COPI notices do not apply in Scotland, approval of PBPP application 2021-0018.
If the processing is of data concerning health or social care, is it for a purpose other than direct care[8]?	Yes	

Consultation

`

[8] direct care: a clinical, social or public health activity concerned with the prevention, investigation and treatment of illness and the alleviation of suffering of individuals. It includes supporting individuals' ability to function and improve their participation in life and society. It includes the assurance of safe and high quality care and treatment through local audit, the management of untoward or adverse incidents, person satisfaction including measurement of outcomes undertaken by one or more registered and regulated health or social care professionals and their team with whom the individual has a legitimate relationship for their care.

Would it be appropriate to seek the views of data subjects or their representatives on the proposed processing?	No
If Y, how will this be done?	
If N, why is this the case?	It is reasonable and in the interest of NHSX as a health organisation to provide data to researchers and technology developers to accelerate the understanding of the COVID-19 disease and support the creation of tools for patient management, diagnosis and risk stratification. Patients' privacy will be protected at every stage and data will be utilised only for the above mentioned purposes.
Would it be helpful to seek advice from independent experts (clinicians, security experts, ethicists etc.) where their specialist knowledge would be useful in understanding and managing privacy risks?	Yes, many experts in the field of radiology and health data management have already expressed their advice on the proposed data processing methodology and are supportive.
If Y, how will this be done?	Already completed and support given
Will any other stakeholder(s) (whether internal or external) need to be consulted about the proposed processing (e.g. NHSE Central team, Public Health England, NHS Digital, the Office for National Statistics)?	Yes, we have consulted experts within NICE, CQC, MHRA, NHSX.
What was/were the outcomes(s) of such consultation?	The experts that we have consulted expressed their support for the proposed processing of patient data.

Datasets and access

Purpose/ process	**Required data items**	**Accessed by (Roles)**	**Storage location**
Data analysis	Email address, password	Clinicians, researchers, technology developers	Data is stored in the AWS cloud infrastructure owned by NHSX. Accepted users will be able to download the data on their local infrastructure

Development of technology solutions (e.g. AI)	Email address, password	Technology developers	Users will do development on their own local infrastructure
Validation of technology solutions	Email address, password	Members of NHSX and AI Lab team, external auditors	Faculty Platform

Data processor[9]

Will the processing be wholly or partly performed on our behalf by a data processor(s)?	Yes.
If Y please give details	1) The Scientific Computing team of **Royal Surrey County Hospital** will be the data collector 2) **Faculty** will provide the Platform environment to manage the data storage (on infrastructure owned by NHSX), and be in charge of some processes for data quality control and data management. Faculty will be using Faculty Platform as a development tool for writing infrastructure code for the warehouse.
Where is the data to be processed by the data processor?	● **Royal Surrey County Hospital** - Data processing takes place in the UK ● **Faculty** - Data processing takes place on NHSX-owned cloud infrastructure based in the UK

If the processing is not completed by a data processor, please ignore the following questions and proceed to the 'Collection of personal data' section …

What assurance has been/will be sought about the/each processor's compliance with the GDPR?	The providers have contractual terms which confirm they are GDPR compliant.
Will the contract use NHS England's standard data processing agreement template?	RSCH Contract Faculty Contract
Will the contract contain standard clauses to require compliance with the GDPR?	Yes, the contract will contain clauses that confirm the provider is GDPR compliant.
Will the contract contain clauses to address the secure transfer of the personal data to a successor data processor should this become necessary or upon the expiry of the term?	N/A

Collection of personal data

Will personal data be collected from the data subject?	No, personal data will be collected from the participating hospitals from the electronic patient record without needing to ask the data subject directly. (COPI Notice and PBPP application 2021-0018 for Scottish data)

[9] 'processor' means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller

Will personal data be obtained from sources other than the subject?	Yes
Will personal data be collected from a third party(ies)?	Yes
If Y, please identify the third party(ies)?	Ethnicity data will also be collected from the segmentation dataset, because it provides a higher level of completeness compared to ethnicity data captured for the purposes of NCCID by research nurses. For sites contributing to both the ISARIC database and NCCID, the clinical variables for the NCCID patients might be extracted from the ISARIC database, instead of requesting the research nurses to collect that information again. Clinical variables and imaging data on Scottish patients will be collected, respectively, from the eDRIS National Safe Haven and the HIC Regional Safe Haven in Scotland.
Is the provision of personal data obligatory or voluntary?	Obligatory under COPI Notice for the English and Welsh data. It is not obligatory for the Scottish data.
If obligatory, why/how is that the case?	Healthcare staff (radiologists, administrative staff) will choose patients whose data, upon pseudonymisation, will be uploaded to the database, with no prior request to the patient. This is to make sure that the most relevant cases are included, to ensure the database is informative and utilisable for its purposes.
What are the possible consequences for a data subject if there is a failure to provide the requested personal data?	There would be no consequences.

Privacy information

How will the data subjects be informed of the processing of personal data about them?	Data subjects would not be directly informed of the processing of personal data about them. Standard Covid Privacy Notice will cover

Accuracy of personal data

How will we ensure the accuracy of the personal data (including their rectification or erasure where necessary)?	There will be input from the healthcare staff involved in the data collection.
How will we monitor the quality of the personal data?	Data quality will be monitored during the collection process: 1) Validation of inputted data is provided by the clinical data templates 2) Additional validation scans are made on the clinical data when it is uploaded 3) Checks for modality and body part are made on the incoming images

	4) Integrity scanners report outliers with regarded to volume or type of data acquired to flag any potential issues The quality and completeness of the data being received by other databases will be guaranteed by the scientific and operations team who lead those databases. Quality checks will also be performed by the NCCID technical team before sharing with users. Additionally, data users will be able to report to NHSX whether there are any issues with data quality.

Subject access and data subjects' rights

How will it be possible to provide a copy of the personal data processed about a particular individual to them (redacted as necessary) should they request access to this information? (If you are purchasing an information management system, you should consider including requirements in the specification about searching and subject access requests.)	No identifiable data held by NHS E
What processes will be put in place to ensure that other data subjects rights can be appropriately applied to the personal data if necessary?	No identifiable data held by NHS E, data subject directed to Data Controller who holds identifiable data

Data sharing (other than between NHSE and NHSI)

Will some or all of the personal data be shared with a third party (other than NHSE / NHSI)	No

If N, please skip outflows in the next section …

If Y, will the personal data be disclosed to a recipient(s) in a country outside the EEA or an international organisation?	

Data flows (including transfers between NHSE and NHSI)

Please supply a data flow map or complete the table:

Inflows					
Sender	*Content*	*Pseudonymised?*	*Mode*	*Security*	*Recipient*
Clinical Sites	Imaging data (X-Rays, CTs, MRs)	Yes. Pseudonymisation takes place on the destination IEP node.	Via IEP	IEP is an established clinical tool for transferring patient images	RSCH
Clinical Sites	Clinical data points	Yes. Pseudonymisation takes place on upload on client-side before submission.	Via NCCID web portal	SSL certificate transfer. User accounts/passwords to access	RSCH
RSCH	Imaging data and clinical data points (including ethnicity data from segmentation dataset)	Yes	AES-256 Encrypted transfer of HTTPS	IAM access control to destination bucket. Audit logs	NHSX data warehouse
RSCH	NHS numbers of patients that need ethnicity assigned by the segmentation team	No. Decryption takes place on client-side	Via secure web portal with dedicated account	NHS numbers are stored encrypted. They are transferred to the client's browser encrypted and decrypted on the fly by the client's browser	Segmentation team

Segmentation dataset	NHS number of patients in NCCID and their ethnicity information	Yes. Pseudonymisation n takes place on upload on client-side before submission.	Via secure web portal with dedicated account	NHS numbers and ethnicity information is uploaded. NHS numbers are hashed on the client's side browser before being submitted	RSCH
HIC Regional Safe Haven	Pseudonymised imaging data from Scotland	Yes	AES-256 Encrypted transfer of HTTPS	IAM access control to destination bucket. Audit logs	NHSX data warehouse
eDRIS/ EPCC National Safe Haven	Pseudonymised clinical data points from Scotland	Yes	AES-256 Encrypted transfer of HTTPS	IAM access control to destination bucket. Audit logs	NHSX data warehouse
ISARIC Data Store in Edinburgh	ISARIC IDs and Pseudonymised clinical data points for NCCID patients, extracted from ISARIC database	Yes	Secure FTP	Technology that encrypts authentication information and data while in transit	RSCH
RSCH	ISARIC IDs of NCCID patients	Yes	Secure FTP	Technology that encrypts authentication information and data while in transit	ISARIC Data Store in Edinburgh

Outflows					
Sender	*Content*	*Pseudonymised*	*Mode*	*Security*	*Recipient*
NHSE data warehouse	Imaging data and clinical data points, including ethnicity information from segmentation dataset	Yes	Accessed via cloud platform	IAM and access control to the platform. Audit logs.	Accepted users of the NCCID database (e.g. software developers and researchers)

Risks

What are the identified risks of the processing? Please complete risk register attached. Available on request, provided separately.

Incident reporting

What plans are in place in relation to the internal reporting of a personal data breach? (NB Unless the personal data breach is unlikely to result in a risk to the rights and freedoms of the individual(s), it will	The warehouse only stores pseudonymised data. The pseudonymised data will be stored on a cloud environment accessible through credentials provided to the teams with data access. All data access is logged and auditable. In any cases of suspicious activity on the cloud storage environment or data

normally need to be reported to the ICO within 72 hours.)	breach, the warehouse administrators will be able to assist NHSX in immediately disabling the credentials of the user in question as the first action, preventing all further access to warehouse by that user. All logs will be retained to permit later examination as required. The warehouse administrators will investigate the nature and scope of the incident in order to apply necessary remediations to the warehouse infrastructure, and communicate this to NHS E without delay. NHS E will be in charge of reporting the breach to the ICO within the required timelines in line with standard IG Policy and procedures.
What plans are in place in relation to the notification of data subjects should there be a personal data breach? (NB Where a personal data breach is likely to result in a high risk to the rights and freedoms of the individual(s), they should be notified as soon as reasonably feasible and provided with any recommendations to mitigate potential adverse effects.)	All access of data stored on Faculty Platform datasets is logged, with logs including the full time and date the data was read. Should there be a personal data breach, Faculty administrators would be able to reconstruct what data was accessed. Faculty administrators would then collaborate with RSCH to put in place processes to retrieve the data subjects, in collaboration with the hospitals the data was collected from. The data subjects identified in this way would be immediately informed by NHS E, and advised on the required remediations in line with standard IG Policy and procedures.

Business continuity planning

How will the personal data be restored in a timely manner in the event of a physical or technical incident?	Network backup and recovery. The data warehouse will be backed up incrementally to a separate S3 one-zone bucket with reduced storage costs. The data could be restored to a new active bucket in the case of incident.

Records Management

Will corporate records be created and / or managed as a result of this processing?	NHS Corporate records will be created and maintained by NHS Covid emergency Hubs and Cells as required
Where will these records be stored?	Record storage within Hubs and Cells
Is there a trained Records and Information Management Coordinator (RIMC) responsible for these records?	Yes, through usual business management

Retention of personal data

What is/are the retention period(s) for the personal data?	For as long as it is required to support research and development in response to the COVID-19 crisis. Data from Scotland can be retained until the 1st Feb 2022 (or 18 months after the approval of PBPP 2021-0018).
What is the basis for this retention period? (Please indicate applicable guidance or rationale)	The database storing personal data will be needed for as long as the NHS will require further support tools to respond to the COVID-19 crisis.
Where personal data are processed outside of NHSE's premises or systems, how will they be securely returned to NHSE for the remainder of the retention period(s) as and when this becomes necessary (e.g. following the closure of the project)?	No personal data will be processed outside of the NHSE's premises or systems. Data processors will not access personal identifiable data, but only pseudonymised data.

Direct marketing[10]

Will any personal data be processed for direct marketing purposes?	No
If Y, please describe how the proposed direct marketing will take place:	N/A

Data portability

`

[10] direct marketing is “the communication (by whatever means) of any advertising or marketing material which is directed to particular individuals” - all promotional material falls within this definition, including material promoting the aims of not-for-profit organisations

Where the processing is based on consent or due to a contract, it is carried out by automated means and the data subject has provided the personal data to us, will it be possible to provide them or a different controller with the personal data in a structured, commonly used and machine-readable format? (NB This does not apply to processing necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller – GDPR Article 6(1)(e).)	N/A

Automated processing

Will the processing result in a decision being made about the data subject solely on the basis of automated processing[11] (including profiling[12])?	No. The technology tools that developers will improve and validate will aim to support the radiologists and healthcare workers do their work better and faster, without removing the human aspect of the decision making process.
If Y, is the decision: ● necessary for entering into, or performance of, a contract between the data subject and a data controller ● authorised by law ● based on the data subject's explicit consent?	N/A
Please describe the logic involved in any automated decision-making.	N/A
Please outline the significance and the envisaged consequences of such processing for the data subject.	N/A

`

[11] examples include the automatic refusal of an online credit application and e-recruiting practices without any human intervention

[12] 'profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's

performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements

ICT

Will we, or the data processor(s), be using a new system to process the personal data?	No: the system exploits existing ICT estate.

If Y to the above question around new systems, please ensure that a System Level Security Policy is completed and risk assessed by ICT before proceeding to the sign off stage below.

Sign Off of DPIA and Processor Checklist

(as appropriate) IG Lead's assessment of

the level of risk

Name	Peter Manser
Signature	
Date	

ICT assessment of

the level of risk

(Where SLSP

provided)

Post	Not Required
Name	
Signature	
Date	

Data

Protecti

on

Officer (Where risk escalated)

Advice of the DPO:	
Name	Not required
Signature	
Date	

Information Commissioner's Office (Where DPIA submitted for review)

Advice of the ICO:
N/A

Caldicott Guardian

(If any or all of the information as part of the initiative, project etc. is subject to the common law duty of confidence – e.g. patient identifiable data)

Advice of the CG:	
Post	Not Required - COPI Reg Notice
Name	
Signature	
Date	

Senior Information Risk Owner (In all cases)

Decision of the SIRO:	
Post	NHS E/I SIRO
Name	Mark Blakeman
Signature	
Date	1/10/2020

Processor Checklist

SUPPLIER INFORMATION
Supplier name: Royal Surrey NHS Foundation Trust
Address: Egerton Road, Guildford, GU2 7XX
Telephone number: +44(0)1483 571122
Name of key contact: Prof Mark Halling-Brown

Telephone number and email address of key contact: +44(0)1483 571122 mhalling-brown@nhs.net	
If your organisation is registered with the ICO please provide Data Protection Notification number: Z486353X	
Is your organisation compliant with the Information Governance Toolkit to Level 2: YES	
SERVICE TO BE SUPPLIED	
Please describe the service which is to be provided: Collection of images and clinical data points to RSCH and transfer to NCCID data warehouse	
CHECKLIST	
1. What data will you be processing on our behalf?	DICOM images (chest X-rays, CTs and MRs) and clinical data points from patients suspected of COVID-19 infection that have undergone a PCR test
2. Will you be processing at an NHS site? If so, where?	Yes, RSCH
3. Will your staff have remote access to NHS England data? If yes, please explain.	No
4. Will you be storing any NHS England data in paper format for any length of time? If yes, how will the data be stored?	No
5. Will you be storing any NHS England data in electronic format?	Yes, pseudonymised DICOM images and clinical data from numerous sites in the UK
6. If yes to Q5: - Do all users of your systems have their own log-in and password?	Yes, access control processes are in place
- How are access rights to systems controlled?	Access to Virtual Machines and storage buckets utilised in this study is administered via IAM in line with our trusts access control policy
- Are back–ups encrypted?	Yes, regular backups are taken to a second location (within our trust) and stored at rest encrypted.
- What protection do you have against malicious code?	All our systems sit within the N3 and behind our trust firewall. All servers have anti-virus scans

	on new files. All unnecessary ports are locked down.
- How often do you apply security patches?	On a regular basis in line with releases from CentOS releases. Patching is tested first on a cloned server.
- How often do you risk assess your security controls?	On a six month basis
- What business continuity/disaster recovery plans do you have in place?	All VMs are cloned and backed up and failure of an entire server room is allowed for through a second redundant server room with failover. All code for collection processes and web portal is versioned and stored in a private bitbucket, hecne collection processes can be restored. All processes are documented on QMS. All databases are backed up on a two hour basis and restore processes are in place.
- Are USB ports/CD writers on staff equipment disabled?	Yes
- Will you be storing data outside the UK? If so, where? What information governance considerations have been taken into account?	No
- Will the data be linked with any other data collections? If so, how will the linkage be achieved?	No
7. What security controls do you have in place for your office premises?	key card access to area, keypad access to offices restricted key card access to data centre
8. What controls do you have in place for the security of your equipment?	All equipment used for this study is in the data centre with restricted key card access, access logs and video surveillance
9. Have you had any security incidents relating to data in the last three years? If so, please explain.	No
10. What policies do you have in relation to information security, data protection and incident reporting? Please provide copies.	RSCH Cyber Security_Policy RSCH Data Protection Policy RSCH Information Security Policy RSCH Password Management Policy RSCH Access control policy Available on request

11. Describe potential disciplinary actions for breach of policy.	HR Disciplinary code under terms of employment
12. What steps do you take to ensure that the people your recruit have the honesty and integrity to handle person identifiable data?	S&M training and regular IG courses and tests
13. How do you ensure that your staff understand the importance of data security and how to keep person identifiable data secure?	S&M training and regular IG courses and tests
14. How frequently do you provide your staff with any training on data security and confidentiality and is their learning tested?	Yearly courses and tests
15. Will you ever transfer NHS England data electronically? If so, how will it be transferred?	RSCH will transfer pseudonymised images and data to the NHSE data warehouse using encrypted transfer utilising the S3 APIs.
16. Will you ever transport any NHS England hard copy data? If so, what security controls will be in place?	No
17. Will you ever destroy any NHS England data? If so, how will this be done and what evidence of the destruction will you provide?	No
18. Will you ever sub-contract work in relation to NHS England data? If so, in what circumstances?	No
Date completed	01/04/2020
Completed by	Prof Mark Halling-Brown
Telephone number and email address	+44(0)1483 571122 mhalling-brown@nhs.net

PM 07Apr2020

SUPPLIER INFORMATION	
Supplier name: Faculty	
Address: 54 Welbeck Street, W1G 9XS	
Telephone number: +44 20 3637 9415	
Name of key contact: Andrew Brookes	
Telephone number and email address of key contact: +44 20 3637 9415 andrew@faculty.ai	
If your organisation is registered with the ICO please provide Data Protection Notification ZA065207	
Is your organisation compliant with the Information Governance Toolkit to Level 2: NO	
SERVICE TO BE SUPPLIED	
Please describe the service which is to be provided: Provision of infrastructure development as well as management and maintenance services for the cloud environment (the “Warehouse”) of the NCCID, and for related validation activities. Faculty will be in charge of some processes for data quality control and data management. Faculty will also create data access accounts when required.	
CHECKLIST	
1. What data will you be processing on our behalf?	Images (chest X-rays, CTs and MRs) and clinical data points from patients suspected of COVID-19 infection that have undergone a PCR test
2. Will you be processing at an NHS site? If so, where?	No
3. Will your staff have remote access to NHS England data? If yes, please explain.	No
4. Will you be storing any NHS England data in paper format for any length of time? If yes, how will the data be stored?	No
5. Will you be storing any NHS England data in electronic format?	Yes, pseudonymised images and pseudonymised clinical data from numerous sites in the UK
6. If yes to Q5:	Yes, access control processes are in place

- Do all users of your systems have their own log-in and password?	
- How are access rights to systems controlled?	All access of data stored in the Warehouse datasets is logged, with logs including the full time and date the data was read. Should there be a personal data breach, Faculty administrators are able to reconstruct what data was accessed. Faculty administrators would then collaborate with RSCH to put in place processes to retrieve the data subjects, in collaboration with the hospitals the data was collected from. The data subjects identified in this way would be immediately informed by NHS E, and advised on the required remediations in line with standard IG Policy and procedures. All user attempts to authenticate with Faculty Platform, whether failed or successful, are logged by the user management and permissions service. Logging includes the name of the user making the authentication attempt, and the Faculty Platform service and project for which the attempt was made. Multiple repeated failed login attempts from the same IP address result in that IP being blocked for a cooldown period.
- Are back–ups encrypted?	The Faculty Platform datasets backend uses Amazon's Simple Storage Service (S3) for object storage. Data stored on the Faculty Platform datasets are encrypted on the server-side using the AES block cipher with a unique 256 bit key. The per file encryption keys are encrypted with a master key which is stored in Amazon's key management infrastructure, and regularly rotated. S3 is configured to refuse uploads of data which does not specify encryption at rest.
- What protection do you have against malicious code?	The Docker containers wrapping application code are inspected by a CVE checking service (https://coreos.com/clair/docs/latest/) to identify vulnerabilities. Libraries used in the development of Faculty Platform are subject to an approval process, including inspection for malicious code and dependencies. The Faculty Platform is audited by a CREST-certified external auditor every six months.

- How often do you apply security patches?	Daily scheduled tasks running on each server hosting Faculty Platform services check for system security patches and install them automatically. Likewise, daily scheduled tasks check for upgrades to the Faculty Platform services themselves and automatically update them.
- How often do you risk assess your security controls?	The Faculty Platform is audited by an external auditor every six months. Additionally, we use an AWS configuration analyzer (https://github.com/nccgroup/ScoutSuite) to verify the integrity of the deployment on a daily basis.
- What business continuity/disaster recovery plans do you have in place?	All VMs are cloned and backed up and failure of an entire server room is allowed for through a second redundant server room with failover. All code for collection processes and web portal is versioned and stored in a private bitbucket, hence collection processes can be restored. All processes are documented on QMS. All databases are backed up on a two hour basis and restore processes are in place.
- Are USB ports/CD writers on staff equipment disabled?	No
- Will you be storing data outside the UK? If so, where? What information governance considerations have been taken into account?	No
- Will the data be linked with any other data collections? If so, how will the linkage be achieved?	Data stored in the Platform might be linked to other data collections, but that linkage would be processed by the other Data Processor, RSCH, as described in the above section
7. What security controls do you have in place for your office premises?	The main office at Welbeck Street can be accessed by dialing a code in a keypad by the entrance door. The code is communicated to employees through a secure password management software and should be updated every 6 months. CCTV cameras are in place on the outside overlooking the entrance and inside the reception area, with appropriate signage to notify this. Camera feeds can be accessed live. The recordings do not include sound and are kept for 31 days. Employees who are assigned a seat in The Office Group space are issued a personal access key card which allows entry to the building at all hours.
8. What controls do you have in place for the security of your equipment?	Employee laptop hard drives are encrypted for anything stored on the hard drive. Employees are mandated to use specific software that can be monitored and serviced through an administrator using device management software called JAMF, which allows the IT team to audit what software is installed, running and turned on.

9. Have you had any security incidents relating to data in the last three years? If so, please explain.	No
10. What policies do you have in relation to information security, data protection and incident reporting? Please provide copies.	Available on request: - Faculty - Security Management Plan - Data Security in Faculty Platform VPC deployments - Faculty Technology Guidebook
11. Describe potential disciplinary actions for breach of policy.	Our usual disciplinary processes are broken down into 3 stages: Stage 1 - Formal Meeting Stage 2 - Second Formal Meeting Stage 3 - Final Formal Meeting If an employee is accused of an act of gross misconduct, they may be suspended from work on full pay, normally for no more than five (5) working days, while the alleged offence is investigated by the People team at Faculty. If, on completion of the investigation and a formal meeting, the disciplinary review panel is satisfied that gross misconduct has occurred, the result will normally be summary dismissal without notice or payment in lieu of notice.
12. What steps do you take to ensure that the people your recruit have the honesty and integrity to handle person identifiable data?	A GDPR course is held for all staff during Onboarding and then once every 4 months. Additionally, a GDPR checklist is filled in at the beginning of every project
13. How do you ensure that your staff understand the importance of data security and how to keep person identifiable data secure?	A data security course is held for all staff during Onboarding and then once every 6 months. Relevant information is contained in the Faculty Technology Guidebook and always available for staff to review.
14. How frequently do you provide your staff with any training on data security and confidentiality and is their learning tested?	Once every 6 months. No tests are carried out.
15. Will you ever transfer NHS England data electronically? If so, how will it be transferred?	No
16. Will you ever transport any NHS England hard copy data? If so, what security controls will be in place?	No

17. Will you ever destroy any NHS England data? If so, how will this be done and what evidence of the destruction will you provide?	No
18. Will you ever sub-contract work in relation to NHS England data? If so, in what circumstances?	No
Date completed	22/07/2020
Completed by	Andrew Brookes
Telephone number and email address	+44 20 3637 9415 andrew@faculty.ai