Research registration form for research studies that collect new data

All research projects collecting Personal Data must be registered with either the UCL Data Protection Office (only available if a UCL researcher is involved) or using this form and submitting to the IFS Data Controller before the data are collected.

This form should be completed if Personal Data is collected and used as part of the research project and, if relevant, before applications for ethics approvals are submitted to UCL Research Ethics Committee (REC). Once the registration is completed, you will receive a Registration Number which can be entered into the Ethics Application Form.

All sections must be completed before submitting this form to the Data Protection Officer. For more information, please see the GDPR information on the intranet and the necessary steps for IFS researchers collecting personal data. Please also refer to the ICO website: https://ico.org.uk/

For definitions, please refer to Annex I.

A. STUDY DETAILS

1	Title of the Study	Enhancing Maternal and Newborn Outcomes: A Comprehensive Evaluation of Obstetric Triage Effectiveness and Midwives Training (THRIVE)
2	Proposed Start Date	26 August 2024
3	Proposed End Date	31 December 2025

B. MAIN CONTACT PERSON

1	Full Name	Britta Augsburg
2	Job Title	Associate Director, Institute for Fiscal Studies
3	Email address	Britta_a@ifs.org.uk

C. PRINCIPAL INVESTIGATOR (If different to main contact person)

1	Full Name	
2	Job Title	

3	Email address	

D. RESEARCHERS (Please list all other internal and external researchers involved in the project)

Name	**Institution**	**Email address**
Antonella Bancalari	Institute for Fiscal Studies	Antonella.bancalari@ifs.org.uk
Julia Loh	Institute for Fiscal Studies	Julia.loh@ifs.org.uk
Prof. Jophus Anamuah-Mensah	Institute of Educational Research and Innovation Studies (IERIS), University of Education, Winneba	jophusam@gmail.com
Dr. Mary Eryam Ashinyo	Quality Lead and Deputy Director, Institutional Care Directorate, Ghana Health Service	keyram1@yahoo.co.uk
Muhammad Shafique Arif	Oxford Policy Management	Shafique.arif@opml.co.uk

E. DETAILS OF PROJECT

Please provide a brief summary of the project, including an explanation of the aims, design, methodology and plans for analysis that you propose to use.

This study aims to understand how to improve the quality of existing health services to positively impact maternal and child health outcomes in Ghana. To this end, we will conduct a rigorous impact evaluation of an innovative training programme to improve obstetric triage, called OTIP (Obstetric Triage Implementation Project). The basic premise of the programme is to empower midwives and make capacity building an integral part of the hospital culture. This will be achieved by selecting seven midwifery champions per hospital, training them in a centrally located hospital, providing them with the necessary materials and asking them to cascade the skills by training their colleagues in their own hospital. We will use quantitative methods to evaluate the effectiveness of the intervention using two complementary approaches. First, we will take advantage of the fact that the trained hospitals were selected based on an arbitrary threshold of deliveries (1,200 per year in 2022) and use a regression discontinuity design (RDD). Second, we will integrate a randomised controlled trial (RCT) into the third and final phase of the national roll-out. Together, these two approaches will provide a comprehensive picture of whether and how the training works. These approaches will utilise both primary sources of data and administrative data from Ghana Health Service. The results of our study will provide important inputs for other parts of the GHS and the Ministry of Health, which are already considering whether and how best to adapt and adopt the approach to streamline other health services they provide.

We further aim to use the study to set the foundation for a longitudinal cohort study in Ghana, with the aim to increase our understanding of child development since birth, particularly focusing on children whose births were triaged as high risk. The proposed longitudinal cohort study will help to document the extent to which such high-risk birth children and their parents might need additional support and will identify the points in time at which these would be most needed. It can further be used to inform what works at different developmental stages to mitigate adverse factors.

Research questions:

1. What is the impact of OTIP on the birth process, maternal and neonatal survival and neonatal health outcomes?
2. How does OTIP affect staff wellbeing and patient satisfaction?
3. What is the role of intervention fidelity and compliance (e.g. knowledge gain and implementation of improved practices, interaction between hospital staff)?
4. How does the nomination of leaders via peers (bottom-up), as opposed to appointed by management (top-down), affect knowledge diffusion?
5. Does top-down versus bottom-up nomination lead to differential selection in leaders' type?

Research methodology:

We will use a cluster randomised controlled trial (RCT) to answer our research questions, where the cluster is a hospital. The RCT will focus on the 25 hospitals in the Central, Greater Accra and Western regions that will receive the OTIP programme during the third and final phase of the national roll-out. Randomisation allows us to create a comparable treatment and control group and thus estimate the causal effect of OTIP. Using a random number generator, half of the 25 hospitals (expected to be 13) will be randomly assigned to receive OTIP training in September 2024, and the other half (12) will receive OTIP approximately 4 months later.

To answer research questions (4) and (5), we cross-randomise hospitals within each treatment arm, assigning half of hospitals within the earlier treated arm and the later treated arm randomly to the top-down nomination arm, and the remaining halves to the bottom-up nomination arm.

We note that the decision on which region of the country to cover in which phase of implementation was not random, which means that the impact estimates are not necessarily generalisable to the first and second phases. To gain a better understanding of the overall impact of the programme, we will use a complementary research strategy, namely a regression discontinuity design using secondary data.

The RDD relies on the fact that the hospitals participating in OTIP were selected on the basis of an arbitrary cut-off. Hospitals with more than 1,200 deliveries in 2022 were considered in-scope, and those with fewer deliveries in that year were considered out-of-scope. The intuition behind the RDD is then that if this eligibility threshold is effectively random, hospitals just above and just below this threshold will be comparable on all characteristics except the outcomes affected by OTIP. This key assumption can be tested by examining how variables unrelated to OTIP (such as accident admissions) behave around the threshold. If this assumption is correct, then any changes in outcomes attributed to OTIP will be manifested as discontinuous jumps at the eligibility threshold.

Analysis plans:

Treatment effects on hospital staff and patient outcomes and satisfaction (research questions 1 and 2) will be estimated using OLS models for continuous outcomes and logistic models for binary outcomes. To increase statistical power, we will combine our RDD and RCT strategies with a difference-in-difference approach whenever before and after data are available. By using data from before and after the implementation of the OTIP training programme, together with eligibility for the programme (either being just above the target threshold or being randomly selected to be treated earlier in Phase 3), this methodology allows causal effects to be estimated in a robust manner. The consistency of the diff-in-diff estimates is subject to the testable assumption that there were no differences in outcomes between treatment groups in the years prior to the introduction of the OTIP Champions programme. We will also use ANCOVA estimators to gain statistical power.

However, estimating the impact of OTIP on these populations is only the first step in our research agenda. In order to design scalable interventions, it is essential to understand the mechanisms by which these impacts are achieved. We are therefore designing the study and data collection in such a way that we can identify factors that are influenced by the interventions and that may represent pathways through which impacts on final outcomes are generated (McKenzie, D., 2012), research question 3. The pathways through which the intervention may operate include increased awareness and adherence to protocols (OTIP training may lead to improved adherence to clinical protocols and guidelines, ensuring standardised and evidence-based maternal and newborn health care practices), and improved performance, communication and collaboration (improved teamwork and communication among hospital staff due to the implementation of OTIP training, leading to more coordinated care for mothers and their newborns). We will measure the impact of the intervention on intermediate outcomes, which we hypothesise are likely mediators of the impact on the primary outcomes, and test whether and to what extent these variables mediate the impact of LM on child outcomes using the Gelbach decomposition method. We acknowledge that this analysis is necessarily suggestive; we will not be able to interpret the results of this analysis as causal, as variation in the mediators could be driven by unobservable factors that are also correlated with child outcomes.

To answer research question 4, we will use OLS models for the continuous outcome variable of knowledge at midwife level, leveraging the cross-randomised nature of the research design. We conduct three types of analysis for research question 5. We first investigate how midwives selected through peer-selection differ from manager-selection by using OLS models to test for the outcomes in which they differ, including network centrality (popularity), personality traits and leadership skills (measured by the puzzle games).

E. DESCRIPTION OF THE DATA YOU PLAN TO COLLECT AND THE INFORMATION FLOWS

Please provide the following information:

- details of any personal data being collected and any special categories of data
- details on who your data subjects are
-the types of data you plan to collect (eg. quantitative, qualitative, testing and assessment data, observation data)
- categories of data collection (e.g. education, income, wealth, demographics, health)

- the methods of data collection
- details of any partners involved in the study, eg other universities or organisation-
-details of any processors being used, eg data storage providers, transcription services, fieldwork agencies
- a description or diagram setting out the information flows, which outlines who will have access to the data at different points in the process-which software and hardware you plan to use
-whether you are planning to merge the data collected with other sources of data?
-a description of the potential risks and mitigation strategies

Personal data being collected

(1) Socio-demographic characteristics, including age, religion, ethnic group, education
(2) Incomes
(3) Contact details
(4) Health data, including birth history

Data subjects

(1) Midwives – midwives chosen as midwife champions for OTIP and randomly selected midwives
(2) Mothers – mothers attending the hospital for postnatal care and in wards who have recently given birth at the hospital in the past 2 months

Types of data

(1) Quantitative survey data
(2) Behavioural games
(3) Observation data – we will organise puzzle games with the midwives and take observations on puzzle scores, leadership skills and team dynamics. For mothers, we will take observations from their child health record books to record their children's health outcomes at birth
(4) Network census data – using online surveys, we will ask midwives to name their friends and colleagues they would ask for help on the job

Categories of data collection

(1) Mothers
    a. Sociodemographic information
    b. Birth history
    c. Delivery experience
    d. Experience at postnatal care
    e. Re-contact information
(2) Midwives
    a. Sociodemographic information
    b. Work conditions
    c. Performance review and incentives
    d. Training
    e. Relations with other healthcare staff
    f. Collaboration and empowerment
    g. Maslach burnout inventory
    h. Personality traits
    i. Knowledge
    j. Contact information
    k. Individual behavioural games
    l. Public good games

m. Puzzle observations

Methods of data collection

(1) Data collection will be done by a team of field enumerators who are trained on all survey instruments. These enumerators will be selected by DataPivot and taken through extensive training. Data will be collected electronically.
(2) Data collection will also be done via online surveys using Qualtrics.

Details of any partners involved in the study, eg other universities or organisation

(1) OPM – Collaborator who handles the contracts and funding
(2) Ghana Health Services – are collaborators without access to PII data
(3) Kybele – are the implementing agency, without access to PII data
(4) University of Cape Coast – we might receive support from UCC researchers going forward. They would not have access to PII data

Details of any processors being used, eg data storage providers, transcription services, fieldwork agencies

(1) CSPro – for collection and storage
(2) Dropbox – for transfer
(3) Cloud – for storage
(4) Qualtrics – for data collection

A description or diagram setting out the information flows, which outlines who will have access to the data at different points in the process-which software and hardware you plan to use

(1) Data will be collected by Data Pivot through CAPI, using CS Pro as well as using Qualtrics.
(2) The collected data will then be stored on encrypted laptops before being transferred by Data Pivot to the UK via encrypted dropbox. IFS will be the recipient.
(3) IFS will move the data from Dropbox to be stored on an encrypted network drive accessible only to the people who need access. PIIs will be removed from the dataset to be used for analysis and the will be accessible to a wider range of individuals. When IFS moves to cloud storage, any data stored in the cloud will be encrypted and based within the UK or EU.
(4) IFS will return some PII data to Data Pivot in Ghana. This will be done via encrypted files on Dropbox.
(5) Data Pivot will continue to store these data on encrypted devices.

We will be merging the data collected with data from hospital midwife rosters.

Potential risks and mitigation strategies

- Low power. Although our power analysis gives us confidence that we have sufficient power, it is based on ICCs from secondary data. There remains a risk that the ICCs are in fact higher and that we will not be able to detect significant effects. We will use the baseline data as soon as they are available to check the assumptions made and will consider shifting resources to interview a larger endline sample if this is deemed necessary. While we cannot increase the size of the clusters, we gain power by increasing the number of observations within a cluster.
- Given the small number of clusters, we are likely to face imbalances at baseline. We will use available secondary data (hospital records) to check for balance in key outcomes (such as maternal and neonatal mortality, composition of hospital staff). If these are unbalanced, we will re-randomise.
- Data quality: We have had extensive discussions with various stakeholders about the use of administrative records. While maternal and neonatal mortality are considered well-reported outcomes, other outcomes are likely to require significant

measurement/data entry. We will therefore collect other key outcomes through our primary data. Primary data collection will be overseen by OPM to ensure quality. We will work closely with the data collection company at all stages of the process, from piloting, through training and field supervision, to data checks. We will also use established protocols for measuring OTIP-related procedures.

- To deal with potential limited availability of hospital staff to complete planned interviews, we will:
    - Provide adequate notice of our visits to hospitals.
    - Allow flexibility in interview timing to accommodate staff schedules.
    - Ensure the opportunity to take breaks if needed.
    - Focus on key questions to minimize interview duration. To this end we will extensively pilot the suggested survey instruments and exclude any superfluous question.
    - Arrive at the interview site in advance to ensure interviews start as planned.
- Presidential elections are scheduled for November/December 2024. Although our data collection is planned outside this period (baseline in August 2024 and endline in January 2025), data collection may still be affected. We have discussed the possibility of collecting data during/just after the elections with GHS and were told that this would not be a problem and that we would be allowed to proceed. However, we will plan for contingencies and discuss the possibility of moving the endline a few weeks into 2025.

F. INFORMATION ABOUT THE PARTICIPANTS

1	Will the study involve children under 18?	Yes – mothers under 18 only if accompanied by guardian over 18
2	Will the study vulnerable participants (other than children)	No
3	How are the participants selected?	Midwives: Midwife champions from nominations and random sample of midwives in the maternity department Mothers: Purposive sampling – mothers attending postnatal clinics and wards at the hospital, if they have given birth at the hospital within the past 2 months
4	How many participants will be involved in the study?	750 midwives, 1250 mothers
5	Where are your participants located?	Hospitals in Ghana

G. PRIVACY IMPACT SCREENING QUESTIONS

	For further information on any of these criteria, please see this document I:\Data\GDPR\Data Protection Registration\background documents\DPIA checklist.pdf and this webpage	Yes/No
	SECTION A	
1	Is this a major project that involves the collection of personal data?	Yes
2	Will the project carry out evaluation or scoring including profiling or predicting?	No
3	Will the project carry out automated decision-making with legal or significant effects?	No
4	Will the project carry out systematic monitoring	No
5	Will the project involve processing of sensitive data or data or a highly personal nature?	Yes, we will process health data
6	Will the project involve processing data on a large scale?	No
7	Will the project involve processing data concerning vulnerable data subjects?	Yes in that we collect some information on new born babies
8	Will the project involve innovative technology or organisational solutions?	No
9	Does the project carry out processing that involves preventing data subjects from exercising a right or using a service or contract?	No
10	Will the project result in you making decisions or treating individuals in ways which can have a significant impact on them?	No

	If the answer is yes to any of the questions in SECTION A please consider carrying out a DPAI.	
	SECTION B	
11	Does the project plan to use systematic and extensive profiling or automated decision-making to make significant decisions about people?	No
12	Will the project process special-category data or criminal-offence data on a large scale?	No
13	Will the project systematically monitor a publicly accessible place on a large scale?	No
14	Will the project use innovative technology in combination with criteria from SECTION A or B.	No
15	Will the project use profiling, automated decision-making or special category data to help make decisions on someone's access to a service, opportunity or benefit?	No
16	Will the project combine, compare or match data from multiple sources?	Yes
17	Will the project process personal data without providing a privacy notice directly to the individual in combination with any other criteria from SECTION A or SECTION B?	No
18	Will the project process personal data in a way that involves tracking individuals' online or offline location or behavior, in combination with any other criteria from SECTION A or SECTION B?	No
19	Will the project process children's personal data for profiling or automated decision-making or for marketing purposes, or offer online services directly to them?	No
20	Will the project process personal data that could result in a risk of physical harm in the event of a security breach?	No
	If you have answered YES to any of the questions in SECTION B, you MUST complete a DPIA.	
	Have you completed a Data Protection Impact Assessment (DPIA)? If not, please document in the box below your reasons why.	

	Reasons for not completing a DPIA (if applicable)	

H. DATA STORAGE AND TRANSFER

1	Will the data be transferred from outside the UK into the UK? If yes, please describe how this will be done	Yes, via encrypted files in Dropbox. The encrypted files will be removed from Dropbox for analysis as soon as possible
2	Will the data be transferred out of the UK from the UK? If yes, please describe how this will be done. If yes, please read the UCL guidelines on International Transfers which can be found here.	Yes, via encrypted files in Dropbox
3	Please list all the locations outside IFS where the data will be stored	Ghana (Data Pivot) potentially OPM, UK
4	Where are the data to be stored within IFS (e.g. standard network, encrypted network, enclave, dropbox)?	Personal data will be stored on an encrypted network drive accessible only to the people who need access. PIIs will be removed from the dataset to be used for analysis and the will be accessible to a wider range of individuals.
5	If your data are to be stored unencrypted at rest, please describe what other security measures will be taken. (e.g data will be de-identified)	Data will be encrypted
5	If your data are to be stored in the cloud and the servers are outside the EU/EEA, please confirm that alternative data protection	When IFS moves to cloud storage, any data stored in the cloud will be encrypted and based within the UK or EU.

	frameworks are in place (e.g. privacy shield).	
6	If your data are to be stored in the cloud, please confirm that you have ready the Information Classification and Handling Policy and the *Data protection and GDPR guidelines for researchers working with personal data* understand the rules regarding storage or data in the cloud.	Yes

I. PARTNERS AND DATA PROCESSORS (Please add additional third parties as necessary)	
NAME OF 1ST THIRD PARTY	
Status (joint controller or processor)	OPM
Location of third party (inside/outside the EU/EEA)	UK
Activity/purpose (e.g. storage, processing, analysis**)**	Collaborator who handles the contracts and funding
Proposed method of data transfer Eg. AES-256 encryption with password.	Data will not be transferred between OPM and IFS
Is there an agreement/ contract in place? For *joint controllers* a GDPR-compliant Data Sharing Agreement is required For *data processors*, a contract with GDPR-compliant clauses is required	PENDING data sharing agreement
If the location is outside the EU/ EEA have you spoken to the Research Services Team about ensuring that the agreement/contract includes Standard Contractual Clauses?	N/A

NAME OF 2nd THIRD PARTY	
Status (joint controller or processor)	DataPivot
Location of third party (inside/outside the EU/EEA)	Ghana
Activity/purpose (e.g. storage, processing, analysis**)**	Data collection
Proposed method of data transfer Eg. AES-256 encryption with password.	AES 256 encryption via dropbox
Is there an agreement/ contract in place? For *joint controllers* a GDPR-compliant Data Sharing Agreement is required For *data processors*, a contract with GDPR-compliant clauses is required	Currently in discussion with OPM what is needed
If the location is outside the EU/ EEA have you spoken to the Research Services Team about ensuring that the agreement/contract includes Standard Contractual Clauses?	Yes

J. LAWFUL BASIS AND PRIVACY NOTICES		
1	Have you written privacy notices to provide to your participants? Please include them with this application.	Yes
2	Are you obtaining (ethical) consent from your participants? Please include the consent forms with this application	Yes
3	What legal basis are you using for processing the data?	legal basis for processing information and in this study is Legitimate Interest

	Have you included your legal basis in your privacy notice in a way that is understandable to participants? Note: this will usually be Article 6(1)(f) *legitimate interests*, where the processing is necessary for your legitimate interests or the legitimate interests of a third party unless there is good reason to protect the individual's personal data which overrides those legitimate interests. In some instances it may be Article 6 (1) (a) *consen*t – if you are unsure, please talk to the Data Protection Officer	(Article 6(1)(f) of the UK GDPR). Our legitimate interest is research into the best way to support women giving birth and the journey of being a parent of a 0–3-year-old.
4	If using legitimate interests as your legal basis, please complete a Legitimate interests Assessment (LIA). A blank form can be found here and previous examples can be found here: I:\Data\GDPR\Legitimate Interest Assessments	
5	If you are collecting Special Category data, please state which Article 9(2) special condition you are using to process these data? Note: This will usually be Article 9(2)(j)-where processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.[1] Have you included your separate condition in your privacy notice in a way that is understandable to participants?	Article 9 (2) (j) Archiving, research and statistics This has been included in the privacy notice
6	How long do you plan to keep your data? Have to added information on data retention to your privacy policy?	We have included the following: "Your information will then be stored for a minimum of 10 years. This is to allow us time to look at the longer-term effects. After 10 years, we will carry out a review to see if there is still useful work that can be done using your data. At any point that

---

[1] In order to rely on this researchers must ensure that processing meets the public interest and that appropriate safeguards and technical measures are in place to protect data during collection, handling, transfer, storage, use and publication.

		we no longer need your data, we will delete it. The data including your personal information that is stored on the IFS secure server will be stored indefinitely to allow for long term follow up. However, IFS will review every five years to decide whether the information could still be helpful in future research. If it is not, then the information will be deleted.”

J. CHECKLIST
(State yes or n/a)

1	Data Protection Impact Assessment	Yes
2	Legitimate Interest Assessment	Yes
3	Consent forms	Yes
4	Privacy notices	yes
5	Contracts with third parties	yes

DATA PROTECTION REGISTRATION

IFS Data Protection Registration Number	**Date issued**
IFSDPR00024	19/11/2024

Annex I - Definitions

Term	**Definition**
Personal data	'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.' [2] If it is not possible to directly identify an individual from that information, then it is necessary to consider whether the individual is still identifiable. You should take into account the information you are processing together with all the means reasonably likely to be used by either you or any other person to identify that individual.
Data subject	A data subject is any person whose personal data is being collected, held or processed
Human research subject	Research is considered to involve human subjects when an investigator conducting research obtains (1) data through intervention or interaction with a living individual, or (2) identifiable private information about a living individual.[3]
Data controller	A 'controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law;
Joint data controller	Where two or more controllers jointly determine the purposes and means of processing the data
Data owner [4]	The individual or organisation who owns the data or who gives permission to access the data. It is possible to be a Data Controller, without being a Data Owner.

---

[2] http://www.privacy-regulation.eu/en/article-4-definitions-GDPR.htm

[3] Definition used by the US Office for Human Research Protections: https://www.hhs.gov/ohrp/

[4] See IFS' *Information Classification and Handling Policy*

Data processor	A ‘processor’ means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller
Processing	‘Processing’ means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction
Sensitive and special category data [5]	Special category data under the GDPR is broadly similar to the concept of sensitive personal data under the 1998 Data Protection Act. Under GDPR, processors must have a specific legal basis for processing special category data. They must still have a lawful basis for processing under Article 6 (*Article 6: Lawfulness of Processing*) in exactly the same way as for any other personal data. Processors will also need to satisfy a specific condition under Article 9 (*Article 9: Processing of special categories of personal data).* Special category data is considered to be more sensitive, and need more protection. Examples of special category data include information about an individual’s: ● race; ● ethnic origin; ● politics; ● religion; ● trade union membership; ● genetics; ● biometrics (where used for ID purposes); ● health; ● sex life; or ● sexual orientation. Although ‘special category data’ is specifically defined in the GDPR, it is important to note that this is not an exhaustive list of the types of data which may be considered sensitive by data subjects, and this may vary based on the context in which the data is being collected and processed.

---

[5] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/special-category-data/

Anonymisation[6]	The GDPR defines anonymisation as '…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'.[7] This means that once the data has been de-identified, it is no longer re-identifiable as the identifiers have been permanently removed. According to the UCL website: *Anonymisation is the process of removing personal identifiers, both direct and indirect, that may lead to an individual being identified.* *An individual may be directly identified from their name, address, postcode, telephone number, photograph or image, or some other unique personal characteristic.* *An individual may be indirectly identifiable when certain information is linked together with other sources of information, including, their place of work, job title, salary, their postcode or even the fact that they have a particular diagnosis or condition.*[8] The GDPR does not apply to anonymised information.
Pseudonymisation	The GDPR defines pseudonymisation as "the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution to an identified or identifiable individual" [9] Pseudonymisation involves removing identifiers from the main dataset and replacing them with codes/unique IDs. The IDs are then stored separately and encrypted. This means that the data is potentially re-identifiable in the future. Unlike anonymised data, pseudonymised data falls within the scope of the GDPR
Data Protection Officer (DPO)	A DPO helps to monitor internal compliance, inform and advise on data protection obligations, provide advice regarding Data Protection Impact Assessments (DPIAs) and act as a contact point for data subjects and the supervisory authority

---

[6] See the ICO's Code of Practice on Anonymisation for further information: https://ico.org.uk/media/1061/anonymisation-code.pdf

[7] See Recital 26: http://www.privacy-regulation.eu/en/recital-26-GDPR.htm

[8] See: https://www.ucl.ac.uk/legal-services/guidance/general-data-protection-regulation-gdpr/gdpr-anonymisation-pseudonymisation

[9] GDPR: Article 4(3b): http://www.privacy-regulation.eu/en/article-4-definitions-GDPR.htm

Legitimate interests assessment (LIA)	Researchers may rely on legitimate interests as the lawful basis for processing personal data. A LIA is three-part test to assess whether legitimate interests applies. It is a type of light-touch risk assessment based on the specific context and circumstances. It will help to ensure that your processing is lawful. Recording the LIA will also help to demonstrate compliance.
Data protection impact assessment (DPIA)	A Data Protection Impact Assessment (DPIA) is a process which helps to identify and minimise the data protection risks of a project. It is necessary to do a DPIA for processing that is likely to result in a high risk to individuals. This includes some specified types of processing. It is also good practice to do a DPIA for any other major project which requires the processing of personal data.