Standard Operating Procedure (SOP)

Processing of personal and study data

Study regions

Version 2.0

1. Aim

This SOP describes the processing of personally identifiable data and the collected study data taking into account the applicable data protection and IT security regulations.

2. Scope of the SOP

The scope refers to all partners of the EUthyroid2 consortium who collect study data from participants. The participating partners are: Bangladesh University of Health Sciences (BUHS), Institute of Marine Research, Norway (IMR), Islamia College Peshawar, Pakistan (ICP), Jagiellonian University, Poland (JU), University of Surrey, UK (Surrey) and Queen's University Belfast, UK (QUB).

3. Abbreviations

Art	Article
BUHS	Bangladesh University of Health Sciences
DIN	Deutsches Institut für Normung e.V. (German Institute for Standardisation)
GDPR	General Data Protection Regulation
gICS	Generic Informed Consent Service
gPAS	Generic Pseudonym Administration Service
ICP	Islamia College Peshawar, Pakistan
IDAT	Personal identifying data
IMR	Institute of Marine Research, Norway
ISO/IEC	International Organisation for Standardisation/International Electrotechnical Commission
JU	Jagiellonian University, Poland
No.	Number
PM	Participant Management
SDAT	Study data (pseudonymised)
SOP	Standard Operating Procedure
SSL	Secure Sockets Layer
Surrey	University of Surrey, UK
TMF	Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (Technology and Methods Platform for Networked Medical Research)

QUB	Queen's University Belfast, UK
UMG	University Medicine Greifswald

4. Types of personal data

The personal data processed includes:

Personal identifying data (IDAT)

(IDAT, e.g. surname, first names, date of birth, gender, addresses, telephone numbers, email addresses): This data is collected directly from the data subjects. They are used to identify and contact individuals.

Date of birth and gender are also used within the study data to check the correct allocation of pseudonymised data, for age calculations and to check the inclusion criteria. For example, only participants within a defined age range and only female participants are included for the study arm of the ambulatory care setting. Information on age (only in the required accuracy, e.g. in whole years or categorisation into age groups) is also used in the scientific analysis.

Administrative data:

The Participant Management (PM) keeps a complete history of contact attempts and contacts (incl. contact result) in the course of the follow-up studies. This serves to support the participants in the course of their participation in the study, both to optimise the re-contacting of the people involved and - in pseudonymised form - to evaluate the response including, if necessary, to improve the response and to determine any necessary adjustments to the recruitment process.

Pseudonymised study data (SDAT):

Survey and health data as well as urine samples are collected and processed according to detailed, standardised protocols. The survey instruments used are self-administered questionnaires and interviews.

The questionnaires are collected via web forms. Interviews are recorded as audio files, transcribed and translated into English.

Evaluation data sets

The scientific analysis of the collected data is carried out by the researchers of the consortium. Researchers outside the consortium have the option of applying for the data obtained. In this case, the data will be anonymised and only the data required for the respective research question will be compiled into a project-related evaluation data set and made available for the research project, taking particular account of any resulting re-identification risk. The application procedure is described on the EUthyroid2 webpage and corresponding forms are available for download.

5. Information, consent and revocation

Participation in the EUthyroid2 study is based on voluntary informed consent (Art. 6 and 7, Recital 33 GDPR, No. 25 and 26 of the Declaration of Helsinki). People invited to participate are informed verbally about the study by the respective healthcare professional and receive written participant information (Part 1 of the informed consents), in which they are informed about the planned study procedures, advantages and disadvantages of their participation in the study and about data protection rights.
The consents are valid for a period of 5 years from the signing of the informed consent form and are extended for a further five-year period in each case, unless a corresponding revocation is received. Within their period of validity, the consents are also valid after the participant's death in the event of death. Consent for the collection of health data, on the other hand, is not automatically renewed after five years. In this case, a new declaration of consent must be submitted by the participant.

The declaration of consent has a modular structure and covers the following areas:

1. Participation in the survey (questionnaires, interview)
2. Data processing and storage
3. Collection, storage and utilisation of urine samples
4. Renewed contact by the study regions (follow-up surveys, invitations to follow-ups, obtaining additional consent)
5. Authorisation to use the data and urine samples for research purposes
6. Instruction on the right to information and cancellation

A participant or his/her legal representative can revoke the consents given at any time and without giving reasons in writing to the regional management responsible for him/her. They can revoke each consent individually or all together (complete revocation).

6. Rights of the participants

Participants are informed in detail about their rights in accordance with the GDPR each time they submit a declaration of consent and have the opportunity to contact the regional management responsible for them at any time with questions.

Participants have the right to free information about the personal data stored about them (Art 15 GDPR) and the right to data portability (Art 20 GDPR). Participants have the right to rectification (Art 16 GDPR) and (selective) erasure of the personal data stored about them (Art 17 GDPR). Furthermore, participants have the right to restrict the processing of their personal data (Art 18 GDPR). Participants have the right to withdraw their consent in part or in full (Art 7 GDPR). Participants have the right to be informed about the use/disclosure of their personal data to or by third parties. To exercise these rights, participants can contact the regional management responsible for them. Participants also have the right to lodge a complaint with the competent data protection supervisory authority.

7. Return of results

Result from the urine analyses will not be returned to donors as single measurements of iodine excretion in spot urine samples are unreliable due to high intra- and inter-day variability.

8. Data access, data transmission and encryption

Personal data may only be stored on computers that require authentication of users (access protection) and access to personal data may only be possible for appropriately authenticated users.

Personally identifiable data (IDAT) is always stored in encrypted form and pseudonymised study data (SDAT) is stored in encrypted form at the locations of long-term data storage wherever technically possible.
The storage of personal data on mobile data carriers (USB sticks, CD/DVD or magnetic tapes from backups) should be avoided and is only permitted if the files are encrypted according to the current state of the art and the password is kept secret (see also 5.1.2).
The transmission of personal identifying data between computers in any network may only take place in encrypted form in accordance with the current state of the art. Pseudonymised study data may only be transmitted via the Internet in encrypted form in accordance with the current state of the art.

9. Pseudonymisation in data collection, storage, processing and use

The storage and processing of study data are always pseudonymised, i.e. without personal identifying data. In accordance with the TMF guidelines on data protection in medical research projects, different, independently generated pseudonyms are used for different areas (domains). The number of people who gain knowledge of these pseudonyms is thus limited as far as possible and unauthorised merging of data sets is prevented.
The pseudonyms for questionnaires and interviews are generated by the UMG, the pseudonyms for the urine samples by IMR and entered into the pseudonym management system gPAS by the UMG. The pseudonyms are assigned to the participants by Regional Management and stored in gPAS. The pseudonyms are random numbers of a defined length that do not allow any conclusions to be drawn about specific persons without knowledge of the assignment list.

10. Duration of data storage

Personal data must be deleted when it is no longer needed to fulfill the specific purposes. Long-term data storage of 10 years is required to fulfill the purpose.
Personal identifying and administrative data of participants are deleted after the end of the follow-up period, when no further direct contact with the participants and no re-identification of individual participants is required. This results in a storage period of 10 years for the personally identifiable data of participants.
The personally identifiable data of non-participants and participants who have completely withdrawn their consent will be stored temporarily for as long as the withdrawal must be documented. This is necessary to prevent the participant concerned from being invited again. As

soon as no new people are included in the study, i.e. after completion of the baseline study at the latest, the IDAT of non-participants will also be deleted.

In connection with the conduct of the study, the participant's documents are also required in paper form. Data on paper must be destroyed in such a way that its contents cannot be reconstructed. Documents in the paper files that are no longer required (barcodes, routing slips, etc.) will be destroyed in accordance with standard ISO/IEC 21964 (DIN 66399) “Office and Data technology – destruction of data carriers” protection class 3 security level 4. Hard discs with personal data of participants will be destroyed in accordance with DIN 66399 protection class 3 security level 3 and CDs/DVDs in accordance with protection class 3 security level 5.

The survey and health data collected will be stored pseudonymised in the EUthyroid2 research database of the UMG. If a participant withdraws their consent to the use of this data, it will be deleted from this database. However, they will not be deleted from the application-specific, second-pseudonymised evaluation datasets already created for scientific evaluations in the context of research projects.

11. Deletion periods

In general, the following deadlines apply to the deletion of data and the destruction of urine samples due to a revocation or the expiry of the validity of the relevant consent:

- Blocking of data and urine samples for further use immediately, where technically possible, otherwise within five working days of receipt of the information or instruction by the respective data-holding institution
- Deletion of interview, examination and health data within one month
- Destruction of urine samples and deletion of the associated data and storage information within three months if this is technically possible. If there are temporary technical difficulties in urine sample storage systems that prevent timely destruction, the urine samples to be destroyed must be effectively anonymised within the deadline. A cancellation of urine samples is thus deemed to have been implemented. The anonymised urine samples will be destroyed within four weeks after the technical difficulties have been resolved.
- Deletion of personal identifying data within one month
- Deletion of pseudonyms and pseudonym assignments immediately: as soon as the data assigned via these pseudonyms has been deleted
- Evaluation datasets that were handed over to researchers as part of the utilisation procedure must be archived for a period of 10 years in accordance with good scientific practice and will therefore not be deleted immediately in the event of revocation. This period of 10 years starts with the publication date of the last publication created with the evaluation dataset.

12. Overarching technical and organisational measures

The processing of personally identifiable data takes place exclusively in the facilities of the study surveys. The study regions are obliged to independently create and maintain the registers of processing activities in compliance with data protection regulations. The following requirements should be met:

1. Separation of personal identifying and study data (personnel, physical, organisational, technical)
2. Online connection of Participant Management to the central database hosted by EUthyroid2 data management of the UMG
3. Data exchange between Ambulatory Care Facility, Regional Management, and EUthyroid2 data management of the UMG
4. Data backup/restore concept for locally stored data
5. Protection of locally processed data against access by unauthorised people (access protection for computers, securing the network, secure handling of mobile data carriers, etc.)

IDAT and SDAT are always processed and stored strictly separately. In organisational terms, this is achieved by separating the study regions into the Participant Management and Regional Management/Ambulatory Care Unit units. Study data is collected, stored and processed exclusively in pseudonymised form.
All data access requires authentication at least using a username and a password. Access rights are assigned depending on the institution from which access is granted and the role that the user has been assigned within this institution. All employees with access to participants' personal data are bound to data secrecy as part of their employment contracts or in accordance with the relevant provisions of the data protection laws of the federal states.
Any electronic data transmission outside of closed networks will only take place in encrypted form after authentication of the sender and recipient by means of certificates.

Joint Controller Agreements are concluded between the consortium partners who jointly process the collected study data, which contain the processes of joint data processing and responsibilities. Data processing agreements are concluded between the Study Regions and the participating Ambulatory Care Facilities.

13. Organisational units with location and tasks

Study regions

Each study region is organised in a **Participant Management (PM)**, a **Regional Management (RM)** and **Ambulatory Care Facilities (ACF)**.

While PM staff may only process personal data such as contact details etc., RM staff only have access to the study data collected.
The healthcare professionals deployed at the Ambulatory Care Facilities represent a special case. The healthcare professionals are not employees of the partners of the consortium but of the cooperating healthcare facilities. They are involved in the recruitment of participants, the collection of consents, the receipt of urine samples and, if necessary, and completed questionnaires. Therefore, certain measures are taken to ensure that personally identifiable data and pseudonymised data are processed separately.
The healthcare professionals receive the following materials for each participant to be recruited:

A) Participant management materials
- Cover letter to the participant
- Informed consent in duplicate with printed Pseudo-ID

- If a digital survey is possible: regional management tablet with certificate stored in the web browser, access data for the digital survey
- In the case of reimbursement of expenses by bank transfer: Forms for the participant's bank details
- Envelope addressed to the participant management

B) Regional Management materials
- Questionnaires with printed questionnaire IDs
- Access data for digital collection of the questionnaires
- Sample container labelled with laboratory IDs
- Envelope addressed to the Regional Management

Documents containing personally identifiable data (e.g. declaration of consent, bank details for reimbursement of expenses, etc.) are enveloped and sealed by the healthcare professionals in an envelope labelled Participant Management before the data is collected. The documents for the study survey are only to be handed over after consent has been given. Paper documents containing study data are enveloped and sealed in a separate envelope labelled Regional Management.

Participant Management (PM)

Tasks of the Participant Management are:

1. Recruitment of the Ambulatory Care Facilities (can be done by PM or DM)
2. Assignment of the pseudo-ID to the questionnaire IDs, interview IDs and laboratory IDs in the pseudonym management system gPAS
3. Dispatch of the study documents to the Ambulatory Care Facilites
4. Recording the paper-based consent forms
5. Management of the contact database and appointments
6. Monitoring the response
7. Contact point for participants and healthcare professionals at the ambulatory care facilities
8. Data protection-compliant archiving of the paper-based consent forms

Ambulatory Care Facilities (ACF)

Tasks of the Ambulatory Care Facilities are:

1. Recruiting participants and arranging appointments
2. Informing the participants and obtaining their consent
3. Handing out the study documents
4. Training of the participants and handing out the intervention materials (for the intervention group)
5. Interim storage of the urine samples under refrigeration
6. Temporary storage of the completed documents for participant management in a sealed envelope addressed to PM

7. Temporary storage of the completed documents for regional management in a sealed envelope addressed to RM

Regional Management (RM)

Tasks of the Regional Management are:

1. Recruitment of the Ambulatory Care Facilities (can be done by RM or PM)
2. Contractual agreement, including order data processing contracts
3. Training of the healthcare professionals
4. Delivery of the study documents
5. Collection of the sealed envelopes for the PM and RM and the urine samples
6. Refrigerated transport of the urine samples to the RM
7. Handing over the envelopes to the PM
8. Processing and shipping of the urine samples according to SOP
9. Conducting process evaluation interviews

Regional data management (R-DM) as sub unit of the RM (applies only for paper-based data collection)

Tasks of the R-DM are:

1. Transfer answers of the paper-based questionnaires to LimeSurvey
2. Store paper-based questionnaires safely (locked and if possible in fireproof- cabinet)

EUthyroid2 Data Management of the UMG

Tasks of the EUthyroid2 Data Management are:

1. Providing the participant management access to the informed consent management system (gICS) and to the pseudonym management system (gPAS)
2. Administration of the informed consents, revocations and identification numbers
3. Inserting analysed data from urine samples into the central data base
4. Production of descriptive reports on a monthly basis, which will provide information on the response rate and the completeness of the completed questionnaires, among other things

EUthyroid2 Laboratory Unit of the IMR

Tasks of the EUthyroid2 Laboratory Unit are:

1. Providing labelled materials for urine samples

2. Providing urine sample transfer agreements
3. Commissioning a courier service that is specialised in worldwide shipment of urine samples in temperature-controlled containers
4. Analysing the urine samples
5. Encrypted Transmission of the analysed data to the EUthyroid2 data management

14. Records of processing activities

According to Art. 30 GDPR, the activities of all processed personal data needs to be documented.

For data processing within the study regions, the templates for the list of activities of the institutions can be used insofar as they fulfill the minimum requirements of Art 30 GDPR.
If required, a sample template can be provided by the Coordinator.

If personal data is processed on behalf of the study region, as is the case of healthcare professionals, for example, the study regions must conclude order processing agreements in accordance with Art 28 GDPR. Art 28(3) lists the aspects that must be contractually regulated. Here too, the study regions can use the templates provided by their institutions.
If required, a sample template can be provided by the Coordinator. No personal data may be processed before the contract is concluded.

If personal data is processed by different institutions with joint responsibilities, joint controller agreements must be concluded in accordance with Art 26 GDPR. No study data may be transferred before the agreement is concluded.

15. Technical and organisational measures

Encryption

The encryption of data, both during storage and data transmission, is an essential component for protecting the personal data of participants. Encryption is carried out according to the current state of the art.

The file system on which the personal identifying data is located must be encrypted. For example, the Encrypting File System on Microsoft Windows (based on the Advanced Encryption Standard) can be used for encryption.

All data connections between the study regions and the UMG are encrypted. In addition, certificates are used to authenticate the UMG servers to the users and vice versa. Access to the UMG servers for the web interfaces of the gICs and the gPAs is only possible with valid SSL client certificates. The personal SSL client certificates are created by the EUthyroid2 data management of the UMG for a restricted group of people. The conditions for applying for an SSL client certificate are that the regional leader provides the full name and official e-mail address of the organisation. The corresponding password is given verbally over the phone. No private e-mail addresses will be accepted.

Once the consents have been collected and digitally recorded in the web application, the participants receive a link to the web-based LimeSurvey survey and a generated access key to complete the questionnaires. Individual access keys are generated for each participant, for each questionnaire. Only the Identification number of the study participants and the corresponding access key are stored in LimeSurvey. Password-protected access to LimeSurvey is regulated via user rights and roles.

The results of the urine samples are transmitted to the UMG by e-mail as an encrypted file, whereby no personally identifiable data is transmitted. The data is read and stored at the UMG in the central database set up for the project.