Data Management Plan for employees at UiT

(English template, version 21.03.2024)

The guidelines for research data management at UiT are found in the Principles and guidelines for research data management at UiT (henceforth: UiT guidelines).

For more information about research data management, see Forskingsdataportalen UiT.

Project details
Responsible for this plan (name): Ellinor Christin Haukland, Siv Gyda Aanes
Date: 30.01.23
Project name and number: HNF1626-22 *Monitoring adverse events caused by systemic anticancer treatment*
Is the project part of larger project? If yes, specify project name and number: No
Affiliation (faculty and institute): Department of Clinical Medicine, Faculty of Health Sciences, UiT – The Arctic University of Norway, Tromsø, Norway The Department of Cancer and Palliative care, Nordland Hospital Trust HF, Norway
Project period: 01.04.23- 01.01.27
- **Ellinor Haukland** (EH). PhD, MD oncology, Department of Oncology NLSH. Postdoc at SHARE-UIS. *Project investigator and main supervisor.* - **Siv Gyda Aanes** (SGA). MD oncology, Department of Oncology, NLSH. *PhD candidate*. - **Siri Wiig**. Professor of Quality and Safety in Healthcare Systems and Centre Director of SHARE-UIS. *Co-supervisor.* - **David W. Bates**. MD, Professor at Harvard Medical School and SHARE-UIS. *Co-supervisor.* - **Carsten Nieder**. MD oncology, Department of Oncology NLSH. Professor at UIT the Arctic University of Norway. *Co-Supervisor* - Gerd Karin Bjørhovde. Dr.Philos, Professor Emerita of English literature. UIT the Artic University of Norway. *User representative.* - Alexander Ringdal (AR), Technical support NLSH. *Project member.* - Tonje Hansen. MD, PhD and Medical chief at NLSH. *Project member.* - Ole Johnny Pettersen, Department of Internal Medicine HSYK. *Project member.* - Hege Sagstuen Haugnes. MD oncology, PhD, Department of Oncology, UNN. *Project member.* - Renate Elenjord. Pharmacist and Head of Research at Sykehusapotek Nord HF. *Project member.*
Short description of the project: Cancer patients experience more frequently patient harm than other patients, mostly from systemic anticancer treatment. To reduce avoidable patient harm from modern anticancer

treatment such as immunotherapy and targeted therapies, precise and clinically relevant measurements for harm and adverse events are needed.

The study aims through three different sub-studies, work packages, to develop electronic methods for measuring automatically injuries in cancer patients, and investigate if electronic patient follow-up (ePRO) can prevent injuries, affect quality of life and survival in patients receiving immunotherapy cancer treatment.

Research has shown that the recommended method for measuring patient harm is not specific or sensitive enough to detect cancer-related patient harm. In recent years, three damage measurement tools for cancer have been validated based on a manual retrospective review of the patient record. Based on existing technology and research, in work package (WP) 1 the aim is to develop and validate a fully automatic damage measurement tool for retrospective detection of damage in cancer patients receiving immunotherapy. The development of the technological solution is done by Datavarehuset Helse Nord and SAS Institute in an already existing solution (Nordic Clinical Analytics Framework - NCAF) which is used daily in all health trusts in Helse Nord, Northern Norway.

In order to prevent avoidable patient harm, it is important to involve the patient in their own treatment. Research on follow-up with electronic patient reported outcome measures reveal more side effects, increased quality of life and survival in cancer patients.
In WP 2, the technology developed in WP 1 is used to prospectively examine whether follow-up with ePRO compared to standard follow-up results in a decrease in adverse events. In WP 3, quality of life and survival are investigated with ePRO follow-up compared to standard follow-up. WP1 is a retrospective diagnostic study. WP 2 and 3 are clinical cohort studies where a prospective intervention group of cancer patients are followed with ePRO during immunotherapy treatment and compared with a retrospective group of cancer patients followed with standard clinical follow-up. WP 2 investigates rates, severity and types of immunotherapy related adverse events in patients monitored with ePRO compared to standard follow-up. WP3 investigates quality of life in the ePRO cohort, and overall survival in patients monitored to standard follow-up.

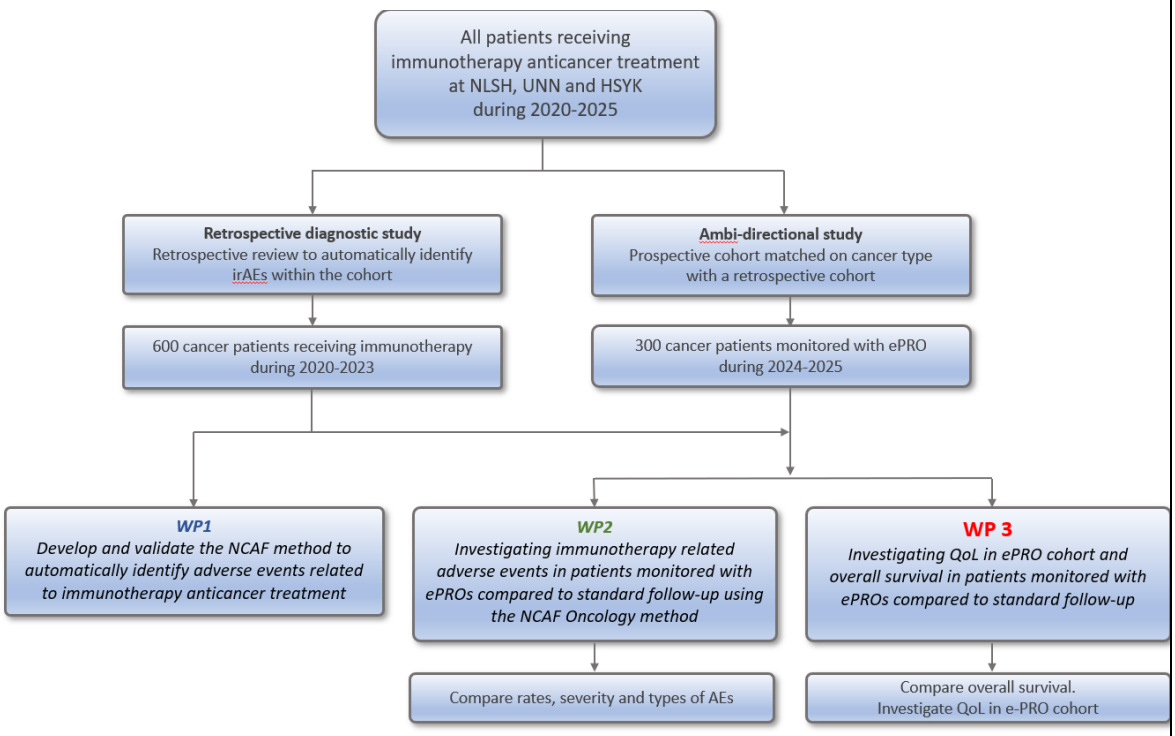
When the study is completed in 2027, the intention is to have developed and validated an automatic measurement tool for the detection of cancer-related injuries for general use nationally and internationally, and new knowledge about how and if digital follow-up with patient-reported outcomes can affect patient harm and adverse events, quality of life and survival.

The study contributes to new knowledge about measurement of patient injuries and electronic follow-up of cancer patients during active treatment with immunotherapy as a new standard of care in our region.

Adverse events are an additional burden for seriously ill cancer patients. By measuring adverse events that occur in cancer patients, we recognize that these events happen and is a part of oncology that need to be taken to account for each individual patient. Knowledge of the occurrence of adverse events in each individual department will be able to say something about status in each department, so measures can be taken at a system level to improve patient safety and the quality of the health service.

The use of personalized electronic follow-up gives the patient the opportunity to participate actively in the course of their illness, regardless of where and who the patient is. Increased involvement of the patient has shown to provide increased empowerment, safety, better quality of life and possible increased survival in other studies. Follow-up with ePRO makes it easier for

healthcare personnel to detect early signs of potential injuries so that measures can be initiated faster to mitigate potential harm.



Funding:	☐ UiT funding: working hours and equipment ☐ UiT funding: direct project funding ☒ External funding: Helse Nord

Responsibilities and rights
Who is responsible for follow-up and revision of this data management plan? Project leader Ellinor Haukland together with PhD candidate Siv Gyda Aanes
Who is responsible for each activity? Project leader Ellinor Haukland together with PhD candidate Siv Gyda Aanes are responsible for collection, documentation and archiving. Mainly Aanes under supervision by Haukland.
How will responsibilities be distributed among possible external collaborators? Will there be a separate agreement on this? Datavarehuset (Helse Nord) and SAS Institute have already responsibility and agreements for existing technology for extraction of electronic patient records through the NCAF solution used by all hospitals in the Helse Nord Trust in Northern Norway. Datavarehuset og SAS Institute are responsible in WP1 for building and the technical testing of possible triggers for patient harm, which then the PhD candidate Aanes will validate. There is an own signed agreement between Kaiku Health and Nordlandssykehuset from 2020 when the ePRO software Kaiku Health where implemented into daily standard use from June 2021. Local DPO, Data protection Officer have assessed and approved the agreement and the software, as well as it has been done Risk and Safety-analysis in 2021 and 2023. Kaiku Health follow GDPR.

Who has the right to manage the data? The PhD candidate, Aanes, main supervisor Haukland, Datavarehuset (Helse Nord), Alexander Ringdal (technical personnel NLSH).
Who can access the data during the project period? These members in the project group: Aanes, Haukland, Nieder, Wiig, Bates, Ringdal, Sagstuen and Hansen.
Who has ownership of the data? Nordlandssykehuset HF

Collecting/generating data
What kind of data will be collected/generated (e.g. observations, simulations, interviews)? What are the sources (e.g. corpora or other raw data)? Qualitative, retrospective and prospective patient data from DIPS (the common electronic health record, EHR, in Helse Nord)
What standards and methods will be used for data collection/generating? There will be done an automatic extraction of patient data from the EHR with help from Datavarehuset for retrospective review of patient journals in WP 1, and these patients in WP 2 and 3 will be retrospectively compared to a prospective group of patients in WP2 and 3 where data in addition will be prospectively extracted from the EHR – collected qualitative and then changed to quantitative before further analysing. All extracted patient data in WP1-3 will be anonymized.
When will the data be collected/generated? Data will be collected from Septmeber 2023 – December 2025
What type of data will be collected/generated (e.g. text, image, numerical data, sound)? Digital qualitative data from the EHR which include age, gender, local hospital, type of cancer and age at diagnosis, cancer treatment, admissions to hospital and treatment under admission, possible adverse events from cancer treatment, radiation therapy under cancer treatment. These will be changed to quantitative data as numerical data for further statistical analyses.
Is there need for extra hardware or software? Yes, in WP 1 and 2 the Nordic Clinical Analysing Framework, NCAF, software from SAS Institute through Datavarehuset and Helse Nord IKT (already in use at Helse Nord) together with the EHR will be used to find possible patient harm/ adverse events. In WP2 and 3 the EHR in Helse Nord in addition with the software Kaiku Health (technical expertise from Elekta Kaiku, Helsinki, Finland) will be needed – already in routine use at Nordlandssykehuset HF. For analysis SPSS will be used as statistical software, available at UIT and Nordlandssykehuset.
Is there need for special expertise for collecting/generating data? There is available special expertise at both SAS Institute and Helse Nord IKT and Datavarehuset to helt with the development of the technical algorithms in WP1, and a own test-area will be developed and used at SAS Visual Analytics platform. There is available statistical help at Nordlandssykehuset with own statistician, Laurent. A.Ringdal at Nordlandssykehuset IKT will help with technical setup and flow of included patients between EHR and the NCAF test-platform in Visual Analytic platform. Idamaria Lehtinin at Elekta Kaiku will help with extraction of data from the Kaiku software for WP3.
Are there any existing data (internal or external) on the topic? If yes, how can they be integrated and reused in the project? **No.**

Documentation and metadata
(See Deposit Guidelines for DataverseNO. Note! This section applies for all kinds of research projects. Good documentation is crucial for your data to be understandable and reusable also in the long term, independently of whether the data will be shared or not.)

How will the data be documented so that they are comprehensible and reusable for yourself and others also in the long term?
(According to best practice, research data should be documented in a ReadMe file which explains column headings, abbreviations etc.)

Data are documented in accordance with the Health Research Act to ensure compliance with the guidelines from the Regional Committees for Medical and Health Research Ethics (REK) regarding confidentiality, sensitive personal information and FAIR principles.

In this project, qualitative data from electronic patient records will be collected and stored in a separate encrypted area for research on the Q-file under the domain of NLSH (own key-folder), where data will be continuously transferred into a main folder for each WP with individual Excel sheets named "Raw data" with an updated date based on editing. All patients will be represented only with NPR-id. Analyses, articles and thesis will be stored in separate folders, named according to which WP the data belongs to.

At the top of the main folder, there will be a ReadMe file for more comprehensive data (including method (title, DOI, contact info, date, location, ownership, financier), method description (protocol, instruments, software), file overview and file specific information with an overview of variables and units. An additional ReadMe file will be created for information that does not fit/right column in Excel. All patients will be anonymised by NPR (a regional specific patient number where you need access to the EHR to identify the patient).

What kind of metadata standard(s) will be used?
(Both open and restricted data have to be provided with metadata according to section 4.6 in the UiT guidelines. Some academic fields have established metadata standards, whereas other fields do not. Examine best practice in your field. See this page for an overview of established metadata standards.)

There is a metadata standard for patient safety, called Patient Safety Ontology, https://fairsharing.org/FAIRsharing.dyj433 .
We will assess wheter this covers the project, but if not, we will use a ReadMe file that covers the areas in the points below, following the standard at UIT`s Open Research Data;

- General background information (title, DOI, contact info, date, location, ownership, financier).
- Method descriptions (protocol, software).
- File overview.
- File-specific information with an overview of variables and units.
- Reference and terms for reuse.
- Descriptions, instructions and protocols for collection, processing and analysing steps.
- Configuration files and log files from processing and analyses.
- Variable list
- Etchical approvals from REK, GDPR for Kaiku Health
- Permissions, licenses, agreements with SAS Institute and Kaiku Health from rights holders

What file format(s) will be used?
Open formats with CSV-files in archive (text format), and numerical and text format in Excel-files. In ReadMe there will be used text format.

What kind of folder structure and filename conventions will be used?
In access-controlled key-folders:
Hierarchical folder structure in relevant categories, with a consistent naming structure on all folders that shows the content of the folders, and folder structure reflecting the file names.
During storage and archiving, the following naming should be used:

- Content in file (e.g patients' data, analysis, article)
- Date/time interval/location
- Name of study/project
- Version nr
- Name/initials of researcher

Is special software for reading/interpreting the data necessary?

In the analysis statistical software will be used. For development of new algorithms in NCAF this software will be needed – but this is already in routine use with other algorithms today. Elekta Kaiku deliver CSV files with variables from extracted data for WP3.

Storage and preservation during the project

What are the procedures for storage and backup, and where will this be done?
NLSH have an own server for safe storage of data (Key). Access is controlled. Head of Research administration at NLSH, Petter Øien, can be contacted for further information about these procedures.

Who is responsible for backup and restoring the data?
Helse Nord IKT is responsible for back-up. NLSH have their own procedure for secure storage

What is the expected file size for the data? 1 TB.

Do you have sufficient storage facilities, or do you need extra services?
Yes, unlimited storage is available, no need for extra services.

If collecting data in the field (out of office), how will the data be safely transferred from the field to the main storage unit?
As an employee at NLSH, one has safe access via mobile office through Citrix or VDI.

Archiving and sharing
According to section 2 in the UiT guidelines, UiT shall as a rule have access to use all research data generated at the institution.

Which data will be preserved, and which will be destroyed at the end of the project?
All collected data are stored for 10 years after the completion of the research project according to the Health Research Act. Identifiable data are destroyed after 10 years. Anonymized result data are further preserved, stored on a secure server.

Will (a selection of) the data be long-term preserved, and how is this decided?

All data are stored for 10 years after the research has concluded, in accordance with the Health Research Act. Identifiable data are destroyed after 10 years. Anonymized quantitative result data are further preserved, stored on a secure server at SIKT, the service provider for the knowledge sector, sikt.no, following their advice on anonymization.
Will the data be made openly available? If only a selection of the data will be openly available, specify which data. Anonymized quantitative result data are further preserved, stored on a secure server at SIKT, the service provider for the knowledge sector, sikt.no, following their advice on anonymization. This will make them searchable and accessible for secondary use in research, educational purposes and for replication and validation. We chose to use SIKT/NSD following advice from UIT regarding personally sensitive data.
If data will not be shared, what is the reason?
Do the data need processing (e.g. conversion to persistent file format(s), depersonalization) before they can be shared? If yes, how will this be done? Yes. Advice from SIKT/NSD will be sought before publishing and archiving data at their service.
Where will data, metadata, documentation and code associated with the data be archived? - Data in the project period: data, metadata, documentation and code stored in safe storage at NLSH in access limited Key-area. - Archiving after end of project: Metadata and anonymized qualitative data will be stored at NSD/sikt.no
What kind of methods or software are needed to get access to the data? Are the methods/software openly available? No, access is ordered by Helse Nord IKT, not open available under the project period. Under archiving access will need to be ordered through SIKT/NSD.
When will the data be made available, and how long will they be stored? Archived data will be available when WP 1-3 are finished, and sufficient anonymization has been achieved. This will be towards the end of the project period, probably 2027. They will be stored for 10 years at NSD/SIKT.
How will the data be licensed for reuse? This will be discussed with SIKT.
Are there other conditions, restrictions or embargo on the use of the data? No

Ethics and consent *(See Research ethics.)*
Does your data include sensitive data? *(If yes, also answer the questions below.)* **Yes**
Are you going to collect informed consent to store and share the data? If so, how? Yes.

For WP1 and the retrospective control group in WP 2 and 3: due to quality assurance and health service research anchored in own institution, consent is not required for collection, preservation or sharing of results after anonymization. An exemption of confidentiality has been granted following an assessment by REK, applications 302945 and 319277. For the prospective cohort in WP2 and 3 informed consent will be collected through a patient information letter, where, if the patient consent, they give their approval digitally in the software Kaiku. The information and study design are granted by REK, application 711238.
How are you going to secure confidentiality and identity protection? The participants are anonymized, where only study personnel with access to secure storage and the EHR can find their identity.