Data Protection Impact Assessment

Version 1.2

Authors: Johan Gustav Bellika, Torje Henriksen, Joseph Steven Hurley, Kassaye Yitbarek Yigzaw, Håvard Blixgård

Date: 2022.06.30

Revision history:

Issue	Details	Who	Date
0.1	Initial version	JGB	2018.06.13
0.2	First draft version	JGB	2018.08.30
0.3	Revised draft based on input from TH and JSH	JGB, TH, JSH	2018.09.04
0.4	Revised draft based on input from KYY.	JGB, KYY	2018.09.06
0.5	Revised document based on input from HB and Guri Rørtveit	JGB, HB	2018.09.10
1.0	Revised document based on input from the privacy protection team at the university hospital of North Norway.	JGB	2018.10.11
1.1	Revised document based on input from the privacy protection team at the university hospital of North Norway.	JGB	2021.01.26
1.2	Minor revision based on revised version of risk assessment version 4.	JGB	2022.06.30

1 Introduction

1.1 Summary

This document contains the data protection impact assessment of the Norwegian primary care practice based research network IT infrastructure, called PraksisNett.

A Data Protection Impact Assessment is a requirement described in Article 35 of GDPR [1] which is introduced by:

> ”Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks.”

This document describes the PraksisNett IT infrastructure in a way that makes it possible for the controllers (the general practice (GP) practices participating in PraksisNett and the institution responsible for the research project that is using PraksisNett) to perform a Data Protection Impact Assessment (DPIA) of their use of the PraksisNett IT infrastructure. A summary of this report will be used for this purpose. This document has been subject to assessment by UNN´s privacy ombudsman and have been approved by his team of security experts and legal advisors.

The document is based on a checklist provided by the Norwegian data inspectorate [2].

1.2 Audience

Who are the readers for this document?

This document is written assuming an audience of general practitioners, clinical research informatics managers, decision makers, researchers with special interest in how the GDPR requirements are supported by the PraksisNett IT infrastructure.

What background knowledge about the research infrastructure do we assume?

We assume that readers of this document is familiar with the PraksisNett data management plan and how reuse of electronic health record data is done using the Snow system [3]. Interested readers can also benefit from understanding how distributed statistical processing is performed using Snow and Emnet systems [4–6].

1.3 Background

1.3.1 Data-flow in the PraksisNett IT infrastructure

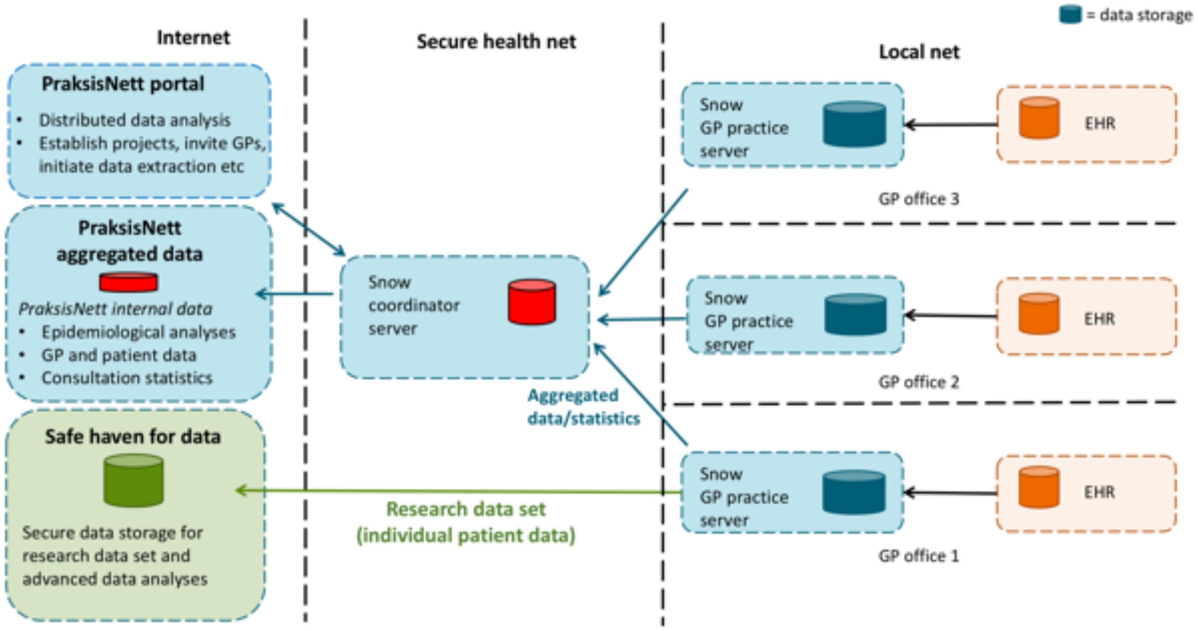


Figure 1. Dataflow in PraksisNett IT infrastructure

The PraksisNett IT infrastructure contains three types of data resources. That is:

1. Research datasets stored inside GP practice servers (blue barrels in Figure 1)
2. Aggregated and anonymous data generated from all participating GP practices (red barrels in Figure 1)
3. Complete pseudonymised research datasets stored inside safe havens (green barrels in Figure 1).

The orange barrels are the EHR database used daily by general practitioners in GP practices to document patient treatment. Before data can be extracted and stored on the Snow GP practice server (in the blue barrels), the GP needs to consent to usage of the EHR data using buypass smart cards (Helse-ID). PraksisNett maintain a protocol of processing approved by each GP according to GDPR article 30. When data is stored on the GP practice server, patients can be identified as potential research subject candidates. Data about both patients and health workers stored in the blue barrels are pseudonymized when leaving the EHR server (black arrows). Pseudonymized research datasets stored inside GP practice servers can only be accessed by personnel with access to the Snow GP practice server. Patients can only be re-identified by the general practitioner or by personnel authorized by the general practitioner, from inside the GP practice.

Based on data in the blue barrels, patient data is first aggregated locally at the GP practice and then aggregated across all participating GP practices. This data is stored inside the red barrel, the report database, at the Snow coordinator server. From there aggregated data can be downloaded using a web service interface to the report database. The PraksisNett aggregated data is a subset of the data stored in the snow coordinator server report database. The flow of aggregated data is shown as blue arrows in Figure 1.

The distributed data analysis client, available from the PraksisNett portal, use the data inside the Snow GP practice servers to produce aggregated data, following the procedure explained above. All results from this processing is stored inside the report database at the snow coordinator server (red barrel) and is available in the distributed data analysis client user interface.

When data collection in a research project (that extract EHR data) is ongoing in the GP practice server, the pseudonymised datasets containing individual patient data will be transferred to the green safe haven. This is shown as the green arrow in Figure 1. The researcher will get access to the complete pseudonymised datasets for advanced data analysis inside the safe haven.

1.3.2 Extraction of patient data from EHR systems

This section is based on the content of [3], which explains in more detail how and what data that can be extracted from EHR systems.

EHR data is extracted from the EHR server using a component we have named "Data reuse component" or DRC for short. This component is run on the EHR server, as shown in Figure 2. The main responsibility of the DRC component is to:

- Replace patient identifiers used in the EHR system with pseudonyms in extracted data.
- Replace general practitioner identifiers in extracted data with pseudonyms.
- Ensure that the same pseudonyms (for each project) are created across collaborating health institutions to allow deduplication of research datasets.
- Ensure that the general practitioner controls whether his patients' data can be extracted and reused.
- Ensure that patients can be exempted from being identified as potential research subjects.
- Ensure that patients can get information about what purposes their EHR information is used for from their GP.
- Ensure that patients can withdraw from research projects they previously have consented to by contacting their GP or the researcher.
- Ensure that patients can get their EHR data, which is stored on the PraksisNett IT infrastructure, exported in a machine readable format.

Figure 2 shows the flow of EHR data locally in a GP practice running a local EHR system. Before EHR data can be extracted and stored on the local data reuse server, GPs must approve use of the EHR data by signing a digital approval of data reuse using their smartcards. When doing so data necessary to fulfil the requirement to maintaining a processing protocol (GDPR article 30) is created and stored in the PraksisNettt project portal. Data about patients are extracted from the EHR system by the DRC running on the EHR server. Before patient data leave the EHR server, identifiers are replaced with pseudonyms. Names, addresses and other identifiers are encrypted before it is transferred to the Snow GP practice servers.

The blue lines shown in Figure 2 shows the flow of pseudonymised EHR data. EHR data is first extracted using the data extraction interface of the DRC. It is then stored as an Export file on the data reuse server, and then imported into the data storage on the Data reuse server.

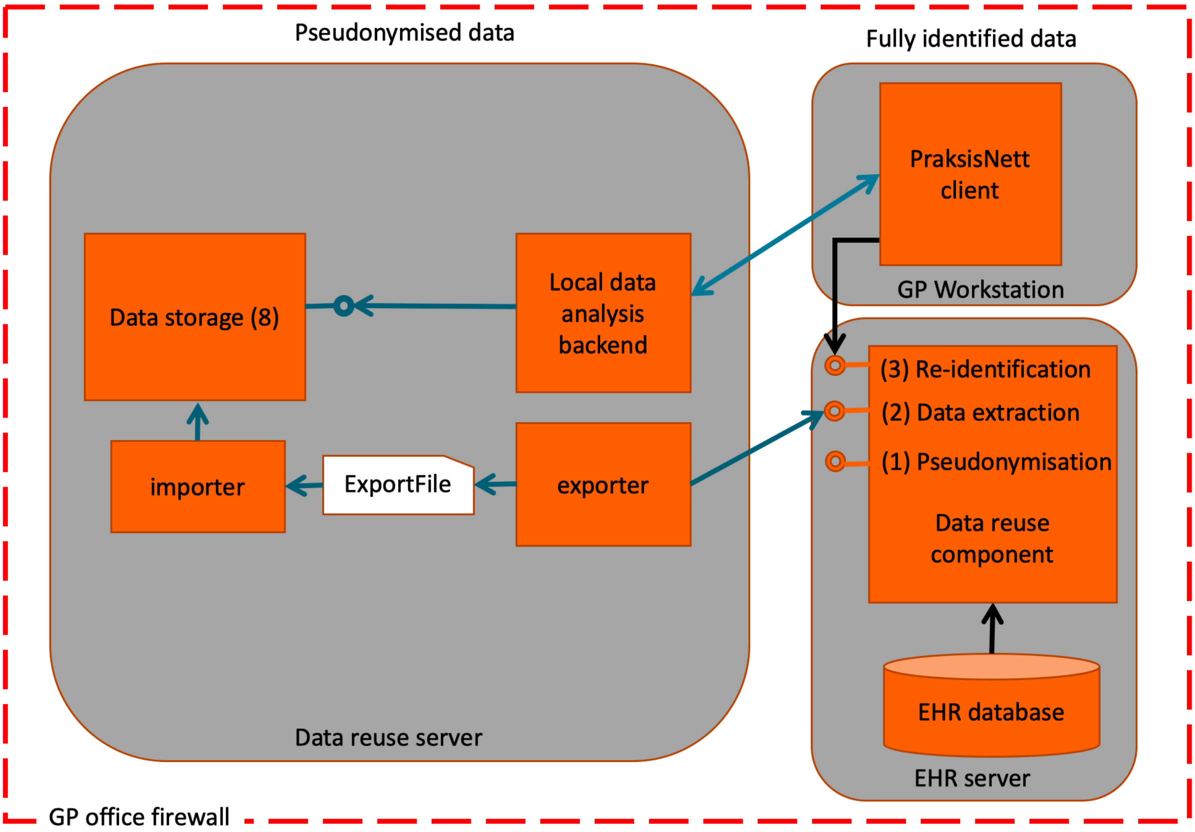


Figure 2. Data extraction mechanism used in PraksisNett IT infrastructure.

When the EHR data is stored in the data storage it can be analysed by the local data analysis backend, which supports the PraksisNett client. The local PraksisNett client can re-identify the GP's patients by calling the Re-identification interface on the DRC. Re-identification is not allowed from the data reuse server. Also, only authenticated and authorized personnel with access to the local area network in the GP practice will be able to authenticate and access the DRC. This is shown as the black arrow in Figure 2.

A research project approved by the PraksisNett board will define the inclusion and exclusion criteria for selecting the patient cohort. These selection criteria will be downloaded to the local data analysis backend. This selection criteria is used to create a local dataset based on the pseudonymised data stored in the Data storage. When a GP has authenticated using the PraksisNett client on his workstation, the list of patients on his patient list will be downloaded to his local workstation from the local data analysis backend. Within the PraksisNett client he can re-identify the patients. This is done by the PraksisNett client making a call to the re-identification interface of the DRC (black arrow in Figure 2). Only GPs who have valid smartcards and authorized by the PraksisNett will be able to access the re-identification interface of the DRC. We assume GPs may want to authorize local medical secretaries and support staff to access the recruitment lists. These personnel will need smartcards for authentication in addition to authorization from PraksisNett to be able to use the PraksisNett client and re-identify patients using the DRC re-identification interface.

2 Data Protection Impact Assessment

According to GDPR a Data Protection Impact assessment should at least contain (GDPR Article 35, 7):

”(a) a systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;

(b) an assessment of the necessity and proportionality of the processing operations in relation to the purposes;

(c) an assessment of the risks to the rights and freedoms of data subjects referred to in paragraph 1; and

(d) the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation taking into account the rights and legitimate interests of data subjects and other persons concerned”

2.1 Systematic description of electronic health record processing operations

The systematic description of processing is split in two parts. This section contains a description of processing done using an electronic health record system. *As this resource is outside PraksisNett responsibility, it is included for completeness.*

How is the information collected?

As described in section 1.3 patient data will be recorded in the electronic health record system used by the GP practice (the orange barrels in Figure 1) as part of medical treatment. The data stored in the electronic health record may also contain information (typically discharge letters) created in other medical institutions like hospitals and emergency wards and transferred to the GP practice electronic health record system using communication software.

How is the information stored?

The patient information is stored in the electronic health record systems. This could be in a GP practice internal server or in servers operated by municipalities or private companies. Some private companies use cloud based storage of patient information.

How is the personal information used?

The patient information is normally used to record medical treatment provided to the patient.

Who has access to the medical information?

In a GP practice, everyone with access to the electronic health record system can normally access all patient information stored in the system. This is necessary to provide medical treatment to the patient. System administration personnel will be able to access the patient information.

Who will be recorded?

Information about patients receiving treatment from the GP.

How can the registered exercise their rights?

Patients can exercise their rights by contacting their GP.

Will there be systematic processing of information about the patients?

Systematic processing of information about the patients may be performed by the electronic health record system.

Are new technology used to process the information?

This is outside our knowledge area as we only have responsibility for processing performed on our data reuse server.

Is the information processed in new ways by old technology that has not been evaluated?

This is outside our knowledge area as we only have responsibility for processing performed on our data reuse server.

2.2 Systematic description of processing in the PraksisNett research infrastructure

This section describes how patient data is processed in the PraksisNett research infrastructure.

Patient data stored in the data reuse server will be used for local statistical processing, for distributed statistical processing and for generating patient recruitment lists during the recruitment phase of research. These purposes will be described below.

Section 1.3.1 and 1.3.2 explains the flow of data in the PraksisNett research infrastructure. All processing performed on the data reuse server will be done on pseudonymised patient data. Processing performed on the GP workstation and the EHR server will be done on patient identifiable data. Processing performed on the coordination server will be done on aggregated anonymous data.

2.2.1 The processing characteristics

How is the information collected?

Information stored in the data reuse server is extracted from the EHR using the DRC component. Additional data (typically questionnaires and measurement data) may also be collected as part of research projects and stored on the data reuse server. See also section 1.3.2.

How is the information stored?

The information is stored on the data reuse server, during the data collection phase of research projects. During the data collection phase of a research project, data may also be transferred to a secure storage (called a "safe haven") where the researcher gets access to it. This is only performed for patients that has consented to participation and processing of the information.

How are the personal information used?

The pseudonymised patient information stored on the data reuse server will be used according to the medical record act paragraph 6 (pasientjournalloven), or in research projects.

Who has access to the medical information?

In the GP practice, only the patient's GP by default has access to the pseudonymised and identifiable patient's information. The GP may also authorise GP practice staff to access pseudonymised and identifiable patient information. When the patient information is transferred to a safe haven, the principal investigator and the ones he authorises will get access to the patient information. System administration personnel in PraksisNett and for the safe haven IT solution will have the possibility to access the information.

Who will be recorded?

This depend on several aspects; first, only patients that belong to a GP that has signed a digital approval of secondary use, using the PraksisNett research infrastructure. Second, for each individual dataset, only patients that satisfy the inclusion and exclusion criteria of data reuse initiatives will be selected. This can be for research purposes or for purposes that follow the medical record act paragraph 6 (pasientjournalloven). Third, only information about patients that have *not* reserved themselves from being identified as potential research subjects will be extracted for research purposes. Fourth, only patients that have consented to the use of their information and participation in a research project will have their data transferred to a safe haven.

How can the registered exercise their rights?

Patients may exercise their rights by contacting their GP or the researcher, depending on how consent for participation is collected. This will vary based on the design of the research project.

Will there be systematic processing of information about the patients?

Yes. We envision systematic processing to support the purposes the patient has consented to, and the processing that follows from the purposes described in the medical record act paragraph 6 (pasientjournalloven). For instance, PraksisNett will produce statistics about the patients, their GPs and the consultations covered by the PraksisNett IT infrastructure. PraksisNett will generate statistics about how many GPs that has approved project participation and how many patients that has provided consent to data processing and participation in the research project.

Is new technology used to process the information?

Yes. The PraksisNett research infrastructure will process the patient information.

Is the information processed in new ways by old technology that has not been evaluated?

No.

2.2.2 The extent of processing

The category of data: All information stored in the EHR system of the GP practice can be reused. In addition, questionnaire data and measurement data collected in relation to research projects will be collected. The patient must provide consent for all use of the EHR data for research purposes.

Number of persons: We envision that the research infrastructure will scale to cover 90 GP practices and 105 % of the Norwegian population when PraksisNett is fully deployed. Additional funding is needed to enable PrasisNett to scale further.

Volume of data: The volume of data is described in [3].

Frequency of processing: Processing of the data will be more or less continuous.

Storage time for personal information: The research datasets will be stored until it has served its purpose for the research project, that is according to the approval of each research project. This can be many years as data may need to be available for validation purposes when the research has been published. A research project may also need to follow patients over several years.

The geographical area: The PraksisNett research infrastructure will cover all health regions of Norway.

2.2.3 The context of processing

In this section, we will examine the processing by looking at internal and external factors that could impact the expectations or the consequences.

What sources of information will be used?

We will use information from the medical records of patients, questionnaire, and measurement data.

What relation exist between the controller and the patient?

The medical information has been collected as part of medical treatment provided by the GP.

What is the power relation between the patient and the controller?

The relation is characterised as a patient – doctor relationship. As the education and experience of a GP within medical subjects normally is much greater than the patient´s, patient would probably accept the recommendation of the GP.

In what degree do the data subjects control the use of their information?

The patient has large control over the use of medical information registered about themselves. When invited to participate in a research study, the clinical information that will be used in the research project will be presented to the patient. Based on this information they can detect errors and assess whether they want to support the research project by giving consent. They will be notified about who will get access to the patient information and where the data will be transferred (to a safe haven). They will also be informed about their right to withdraw consent provided to processing of the information about themselves.

Describe how the processing will be perceived from the patients point of view. Can the processing be seen as unpredictable by the patient?

As the processing is untraditional, the processing may be viewed as unpredictable by the patient. Efforts will be made to clarify how processing of the patient's information will be performed. This will be the task of each individual researcher when formulating the study invitation letter and consent forms. The management board of PraksisNett will ensure that sufficient information and clear information will be provided to the patient during the evaluation of the research proposal.

Do the registered have an expectation of confidentiality?

Yes. Patients have an expectation of confidentiality.

Do the registered have an expectation of correct information and a necessity of processing?

The patients expect that information about them is correct. Information extracted from the EHR for use in the research project will be provided to the patient according to the requirements of GDPR. We believe patients understand the necessity of processing the information for the purposes described in the medical record act and for research.

Do the registered expect privacy?

Yes.

Will the processing include vulnerable persons as children and patients?

Yes.

Does experience from similar processing exist?

No.

Describe potential progress within technology or security.

The innovation exploited by the PraksisNett research infrastructure is that a privacy preserving system can be used to do statistical computations and locate potential research subject based on the local data storage within GP practices. Only their regular doctor, the author of the clinical information, need to see their records to have the opportunity to participate in relevant research studies.

Does examples of worried voices exist about this method of processing health information?

No, we have not heard such worried voices about this way of processing patient information yet.

Will we process personal information from different sources, collected for different purposes and from different controllers?

Yes. We will also get access to medical information sent to a GP practice from collaborating health institutions. This will typically be discharge letters from various health institutions. Research project may also collect additional data, like medical measurements and questionnaire data which may be stored on the data reuse server.

Will data from several sources be combined to produce new information about the registered?

Yes. We envision combining medical data from several sources. Information from other sources that is stored in the local EHR are discharge letters, test results and evaluations performed by other medical services. It is also possible to envision that data about the patient from several types of health institutions may be combined. For example, data from the patients themselves, their medical sensors and activity trackers, data from emergency wards, nursing homes, home care, physiotherapy, hospitals etc. This will improve our ability to get a more correct and complete “view” of the patient’s condition. However, using information from all these institution is not likely to be possible in the near future. However, to ensure validity of this DPIA document beyond the current plans, this vision is included to ensure future validity of this DPIA document.

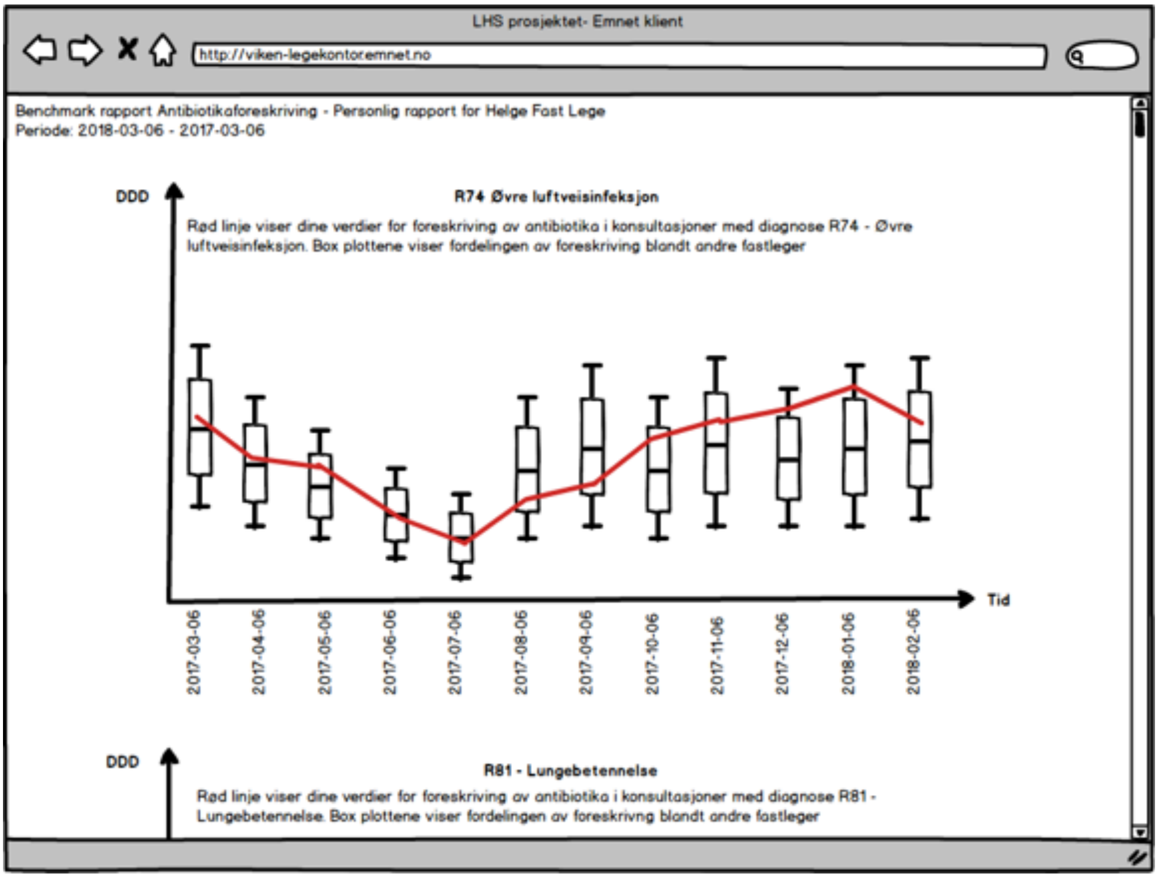
2.2.4 Prerequisites for processing of patient information on the data reuse server

Before processing of a patient’s EHR information using the PraksisNett infrastructure become possible, a digital approval of participation in a project need to be made by the GP responsible for the patient. When such an approval exists, data can be extracted by the DRC and stored on the GP practice data reuse server. Common to all uses is that a purpose needs to be registered in relation to every dataset that the GP creates. For a research project the purpose will be research. However, every research project must specify its objectives beyond being “research”.

2.2.5 Local statistical processing of patient data on the data reuse server

When a GP has created a digital approval for patient data usage, he can use the local PraksisNett client to create working lists for patient administration, internal control and quality assurance. These working lists can be used for providing health services to patients according to treatment guidelines. Examples include, making a working list of diabetic patients for regular check-ups, making a working list for elderly that need a flu shot in the influenza season.

2.2.6 Distributed statistical processing of patient data on the data reuse server



Benchmark rapport Antibiotikaforeskriving - Personlig rapport for Helge Fast Lege
Periode: 2018-03-06 - 2017-03-06

R74 Øvre luftveisinfeksjon

Rød linje viser dine verdier for foreskriving av antibiotika i konsultasjoner med diagnose R74 - Øvre luftveisinfeksjon. Box plottene viser fordelingen av foreskriving blandt andre fastleger

R81 - Lungebetennelse

Rød linje viser dine verdier for foreskriving av antibiotika i konsultasjoner med diagnose R81 - Lungebetennelse. Box plottene viser fordelingen av foreskrivng blandt andre fastleger

Figure 3. Individual feedback to GPs on atibiotics perscriptions

Distributed statistical processing of patient data is performed by first aggregating data locally at each participating GP practice. This aggregated data is then aggregated again across participating GP practices (minimum 3 GP practices have to participate) to produce the overall statistical results.

Distributed statistical processing is necessary to perform tasks like research, quality assurance, quality improvement work and disease surveillance. This mechanism makes it possible for GPs to do quality assurance and quality improvement work by comparing their practice to their peers. By having this possibility, they can also participate in quality improvement work and discuss aspects of their practice with their peers. The purpose of this processing is improved medical services for the patients.

An example of such usage is individual GP feedback on their antibiotics prescription practice as shown in Figure 3. The aggregated level (the boxplots) shows the average prescription rates and the distribution across participating GPs combined with the individual prescription rates (the red line).

Similar feedback can be generated for many different areas such as sick leave, referral practice, microbiology testing, drug prescribing in specific medical areas.

2.2.7 Generating patient recruitment list from patient data

For usage of patient data for research purposes, consent from the patient is required, both for processing the patient information and for participation in the research project, unless an exempt is approved by the regional ethics committee and accepted by the PraksisNett board. Every research project therefore needs to create a recruitment list of patients and send an invitation letter to each potential research subject. The invitation letter text will be prepared by the researcher and must satisfy the requirements from GDPR for informed consent. These requirements are described in article 13, 14 and 15.

The researcher will also define the inclusion and exclusion criteria that will be used to create the recruitment list of potential study participants for each GP that has consented to participate in the study. The GP, which has the responsibility for the potential study participants, will use this recruitment list and the prepared invitation letter and consent forms to create individual study invitation letters to patients. The GP may authorise GP practice support staff to access the patient information and the invitation letter. Information about who has accessed the patient information, who will be accessing the patient information in the future, and where patient information will be transferred after the completion of data collection, needs to be included in the invitation letter and the consent forms. The patient must also be informed about their right to withdraw their consent at any time, their right to get a machine readable copy of the patient data, their right to get a list of persons or category of personnel that have accessed the patient information, and their right to get their data corrected in or deleted from the study dataset.

2.3 Purposes of the processing (the legitimate interest pursued)

What is the purposes of the processing?

The purposes of the processing are: research, patient administration, internal control, quality assurance, quality improvement work and disease surveillance. Every research project also have specific research objectives.

The primary purpose of research is described in the Helsinki declaration article 6 [7] which is:

> "The primary purpose of medical research involving human subjects is to understand the causes, development and effects of diseases and improve preventive, diagnostic and therapeutic interventions (methods, procedures and treatments). Even the best proven interventions must be evaluated continually through research for their safety, effectiveness, efficiency, accessibility and quality."

In addition to usage of patient information for research purposes, patient information can be used for the purposes described in the medical record act (pasientjournalloven) paragraph 6, which is administration, internal control, quality assurance of medical treatment. Usage of patient information for these purposes enables improved medical treatment of the patients and ensuring that each GP is able to compare his practice to the practice of his peers. Then the GP will have a tool to know whether his practice is within or outside the normal ranges of medical practice. The example shown in Figure 3 shows the "normal practice" as a boxplot visualising the median, 25 and 75 percentiles (the box) in addition to the calculated minimum and maximum.

Paragraph 6 also states that processing for internal control and quality assurance should as far as possible be done without using the name or national personal id number (personnummer) of the patient which is satisfied by the method used in our case.

All purposes listed under paragraph 6 of the Medical record act (Pasientjournalloven), are legal purposes that the Patient cannot object to. Also, processing in relation to asking for consent for usage of the patient information, is within legal uses of the patient information.

Also, as disease surveillance is a task aimed at protecting all citizens and is covered by a specific law that applies to all levels of the health service, the patient cannot deny processing of their information for this purpose.

Will the information be used for control purposes (such as tax, custom, police or insurance)?

No. The primary purpose of PraksisNett research infrastructure is medical research.

Could the purpose be to make decision about individual persons based on systematic and extensive analysis of personal aspects?

Yes, such usage could be envisioned. However, only after a consent from the patient. For example, patients will be invited to a research project based on a match with the inclusion and exclusion criteria specified by a researcher.

Will processing of personal information be aimed at decisions that will influence the registered?

Yes, the intention of research is to improve medical treatment, which will have an influence on patients´ health.

Will the information be used to profile the registered?

Yes. The purpose of the research infrastructure is to identify patients that fit the patient inclusion and exclusion criteria of research projects. Also, each GP may wish to find a certain group of patients (like all diabetics) that he wants to follow up on a regular basis.

Will the personal information be used to uncover unknown traits or to discover pattern of the registered person?

Yes, such usage could be envisioned, but with consent from the patient. For example, the specific objective for a research project could be to uncover unknown traits and patterns among a group of patients.

Will the information be processed for other purposes?

That would be tertial purposes. No, not without additional consent from the patient. The PraksisNett management board will ensure that all objectives are listed in the consent forms and invitation letters to the patients.

2.4 Sources, recipients, information security and responsibilities

Are all controllers, processers and potential processors identified?

As the PraksisNett IT infrastructure is currently under construction it is difficult to provide a complete list of controllers and processors. However, Figure 1 provides an overview of the type of institutions that will be controllers of the information about patients. Several potential processors are also identified.

These controllers are:

GP practices: The PraksisNett research infrastructure will recruit 90 GP practices or more to participate in the research infrastructure. Each GP practice will be given a Snow appliance box (a data reuse server) where pseudonymised patient data will be stored. The Snow appliance box may also store questionnaire data and measurement data collected in the GP practices.

Safe havens: When the data collection phase of a research project, (that use EHR data extraction), has started, data belonging to each individual research project will be transferred to a safe haven. The partial datasets, consisting of the pseudonymised records belonging to the patients that have consented to participation, will be transferred directly to the safe haven. The data will be encrypted during transfer. Two such service currently exists [8, 9].

Research institutions (if GDPR Article 26 apply for the research datasets on the Snow appliance box): The GP office and the research institutions are joint controllers for the research dataset stored on the Snow appliance box. These datasets may eventually be transferred to a safe haven, as explained above. Being joint controllers influence what information patients receive when invited to participate in a research project. The division of responsibilities between the GP office and the researchers will be clarified as part of the agreement between the researchers and the GP office when GPs accept being part of a research project.

Processers:

University hospital of North Norway: All GP practices will need to sign a data processing agreement (Databehandleravtale) with UNN which has the responsibility for the daily operation of the IT infrastructure. All Snow appliance boxes will be operated using a centralised operation system by the Snow team. Members of the Snow team will be able to remotely log on to each Snow appliance box to do system maintenance. Operation of all the Snow appliance boxes will be logged to a remote and centralised monitoring systems operated by the Snow team. From the Snow appliance box, members of the Snow team will normally not be able to log on to the EHR server and do system installation and maintenance of the Data Reuse Component (DRC). Employees of Medrave Software AS will be able to log on to EHR server and do system maintenance of the Data Reuse Component. UNN can decide to transfer the responsibility for operation of the infrastructure to a third party at any time. Such a transfer will be governed by an agreement that satisfies the requirements from the «The Code of Conduct for information security in the healthcare and care services" and GDPR (Article 28.3).

Medrave software AB: As part of the research infrastructure the Medrave tool will normally be provided to the GP practice. The Medrave tool will extract pseudonymised patient data from the EHR system and use the information to provide feedback to the GPs in the GP practice as part of the quality improvement processes. Personnel from Medrave Software AB will be able to log on to the EHR servers and do system maintenance on the Medrave installation. Medrave Software AB has signed a third-party agreement with UNN according to the requirements in the «The Code of Conduct for information security in the healthcare and care services" and GDPR (Article 28.3). This agreement allows personnel in Medrave Software AB to install software and log on to any Snow appliance box in PraksisNett. However, in practice this option has not been necessary.

Safe havens: When research datasets are transferred from the Snow appliance boxes to the safe haven environment, the responsibility of the PraksisNett ends. Agreements will be made with the safe haven responsible organisations before research datasets can be transferred there. Only data about patients that has consented to transfer of their data will be transferred.

Are all recipients of the patient information identified and documented? (employees, data processors, external entities etc)

Two types of patient information will be processed in the PraksisNett IT infrastructure; pseudonymised information and fully identified patient information. By default all employees of a GP practice normally have access to fully identified patient information stored in the GP practice EHR system. The PraksisNett IT infrastructure will become a new channel for accessing the patient information.

Fully identified patient information:

Such access is only possible if a number of conditions have been satisfied. These are:

- The GP can authenticate against HelseID (must satisfy a number of conditions to achieve this access)
- The GP has approved reuse of his patients´ EHR data
- The GP is authorised by the PraksisNett staff to access the re-identification interface of the DRC from his workstation.

If all above conditions are met, the GP of a patient by default has access to the fully identified record of the patient in a research dataset. The GP may authorise GP practice staff to access the fully identified study research dataset information and the invitation letter, data processing consent and study participation consent form and clinical information overview that will be mailed to the patient as part of the recruitment phase of a research project.

Pseudonymised patient information:

By default, all GPs that have approved participation in a research project (by making a digital approval) in a GP practice will be able to access the full pseudonymised research dataset stored in the data reuse server for the GP practice. GP practise staff may also be authorised by the GP to access the pseudonymised dataset.

How is the patient information shared internally in the practice?

Fully identified patient data can only be accessed by the patient's GP, and the staff that have been authorised by the GP. Pseudonymised patient information may be accessed by all GPs that have approved participation in a research project or data reuse initiative (see more details in [3]).

Which information is shared with whom?

Datasets are created based on a specification of variables available for extraction from the EHR system (see more details in [3]). Based on this specification a dataset matching the inclusion and exclusion criteria will be created. Before the dataset is created the GP needs to approve including his patients´ records in the dataset. The information in the dataset will be shared according to the description above for internal users and PraksisNett study support staff.

What is the purpose of sharing access to the patient information?

The purpose of creating the dataset can be patient administration, internal control, quality improvement work, disease surveillance or research.

What information is shared externally, for which purposes and on what legal basis?

Only datasets created for research purposes will be shared externally. In such cases, the pseudonymised dataset is transferred to a safe haven, where it is stored until deleted.

System administration personnel in the Snow team, responsible for operation of the IT infrastructure, may visualise patient information stored on the data reuse server (pseudonymised and encrypted data) as part of system maintenance through remote desktop solutions.

Is the information shared with countries outside EU/EØS area on what legal basis?

No, not according to current plans.

Will personal information be transferred to third states or international organisations?

No.

Describe what precautions have been taken to protect personal information (confidentiality agreements, data processing agreements, norms for information security, security measures).

- All personnel responsible for system operation of the PraksisNett IT infrastructure have or will sign a confidentiality agreement.
- All practices that join the PraksisNett IT infrastructure will sign a data processing agreement (databehandleravtale) with UNN, which is responsible for the daily operation of the PraksisNett IT infrastructure. UNN has its own privacy ombudsman, extensive experience within medical research, handling patient information and evaluating information security risks.
- **If Article 26 apply:** All research institutions which act as joint data controllers with the GP offices will sign a data processing agreement (databehandleravtale) with UNN.
- The new Norwegian version of «The Code of Conduct for information security in the healthcare and care services" (which is adapted to GDPR) will be followed.
- Many security measures have been taken to reduce the risk of security breaches to a minimum. They are all described in the "*Snow team information security strategy*" document. Access to this document is restricted. See also list of additional security measure below.
- The architecture of the PraksisNett IT infrastructure minimize the risk of information security breaches by design.
- 10 risk assessment processes have been performed so far, both internal and external, covering different aspects of the PraksisNett IT infrastructure.
- Risk assessments covering the research infrastructure will be revised every year.
- Third party processors, like Medrave software AB, which may install software on the Snow appliance boxes to process patient information, have to sign an agreement with UNN which states the rights, duties and responsibilities with regard to information security, GDPR (article 28, 3) and the «The Code of Conduct for information security in the healthcare and care services".

Additional security measures:

- Each data reuse server is equipped with an on/off switch. The GP practice may stop processing of patient information by powering off the data reuse server, pulling out the power supply or disconnecting the network cable. All these actions will effectively stop the processing of patient information.
- Pseudonymisation. To reduce the risk of exposing identifiable patient information, all data on the data reuse server is pseudonymised. Identifying information, such as names, addresses, etc are stored encrypted on the data reuse server. The data reuse server will not contain any information that can be used to directly identity patients or health workers. This kind of information will be stored on the EHR server.

- Physical protection of the data reuse server. The data reuse server must be physically protected in the same manner as the EHR server.
- Centralized monitoring. All activities on the data reuse server are automatically logged to a centralized monitoring system. Persons that attempt to legally or illegally access the data reuse server are logged to the centralized monitoring system. The centralized monitoring system will regularly produce reports about the activity on the data reuse server. Illegal attempt to access the data reuse server will become visible in these reports. Routines for examining the reports should be established by the GP practices. Then illegal attempts to access the information will likely be discovered.
- Minimizing EHR data extraction. The data reuse server will try to minimize data extraction by extracting EHR data only once per day, preferably in the evening / night, when no one is using the EHR solution. This will minimize the experienced load on the EHR server.
- DRC-software. We have chosen to install the DRC on the EHR server. This minimize the likelihood of exposing patient identifiable information about patients and health workers and satisfy the requirement from GDPR in the best possible way. However, the solution introduces the possibility of affecting the stability of the EHR server. A number of measures to minimize the risk of affecting the stability of the EHR server is established.
- Blocking information access from the data reuse server. To eliminate the possibility of getting access to patient identifiable patient information via the data reuse server, such access is blocked.
- Logging of activity on the data reuse server. To expose internal access to patient information all access to datasets will be logged. The activity will be reported regularly as part of the information security report to the GP practice. Legal and illegal access to patient information will become visible in these reports. All system administration personnel will use personal login accounts. All system administration will be logged with a reason for login and will also be reported as part of regular reports to the GP practice.
- Transfer of research datasets. To reduce the risk of losing research datasets (physical loss), datasets will not be transferred to the researchers. The research datasets will be transferred to safe havens where the researchers get access to it in a secure environment.
- Other security measures. A number of other measures to ensure health workers access to EHR information and reduce the risk of exposing EHR information about the patients and the health workers is used. One of the most important ones is following «The Code of Conduct for information security in the healthcare and care services".

Do you have an agreement or contract with external organisations about the mutual understanding and responsibilities?

Yes. All GP practices running a Snow appliance box must sign a contract with PraksisNett that clarifies rights and obligations for participation in PraksisNett.

The GP practice will also sing a data processing agreement with the University hospital of North Norway that clarify the responsibilities of both parties.

Do the agreement reflect the limitations for sharing personal information?

Yes.

The relation to processors, for each processor:

Are all processors identified and is the relation to each one clarified though agreement (article 28,3)?

All planned processors have been identified. These are the Snow team at UNN which is responsible for daily operation of the PraksisNett IT infrastructure, the safe haven institutions (TSD at UiO and SAFE at UiB), the responsible research institutions that the researchers belong to, and third-party vendors like Medrave software AS providing IT support for the quality improvement software tool and the DRC installation.

Are the registered right and freedom ensured in the agreements?

Agreements between the PraksisNett consortia and each researcher will need to be made to ensure that the rights and freedom of the registered is ensured. The patient has the right to withdraw his consent for usage of his information in a research project. While the data is stored inside the data reuse server of the GP practice, the GP himself can delete the record in the research dataset using the PraksisNett client. An access log for each dataset (corresponding to a project) is kept at the data reuse server. When the research dataset has left the PraksisNett research infrastructure, the researcher must ensure that patients that withdraw from a study are deleted from the research dataset.

To withdraw from a study, the patient needs to contact his GP or the researcher. One of these two will identify the patient's pseudonym, depending on how consent to participation is collected and whether the patient is identifiable or not to the researcher.

The GP and the researcher will instruct each other about deleting the record corresponding to the patient pseudonym in the research dataset. A satisfactory access log also needs to be kept by the safe haven. The access log may be provided in case a patient wants to know who has accessed his pseudonymised data.

When the GP and the research institutions are joint controllers the patient will be informed about what information that is collected by the GP, what information is collected by the researcher and others (support article 13 and 14). The patient will be informed that the GP and the research institution are joint controllers. The patient will be informed about who is the privacy ombudsman for both data controllers and the contacts points with regard to executing their rights. The patient will also be informed about the division of responsibilities between the data controllers.

Are the privacy principles, for instance limitation of purposes, data minimisation, storage etc. handled in the agreement?

Not to a great extent, as it would be too extensive to be described in the agreements. However, the rights of the patients are ensured by the architecture and the functionality of the PraksisNett IT infrastructure (limitation of purpose, data minimisation, data correctness, logging of use, safe storage, ensuing the opportunity to get the data deleted from the dataset, the right to get a copy of the data in a machine-readable format, etc). The agreements state that only patients that have consented to participate in a research project can have their information exported to a safe haven.

How is the information you share with a recipient secured at the receiving end?

Research datasets are transferred from GP practices to safe havens (TSD at UiO and SAFE at UiB). As part of an agreement with the receiving safe haven, a description of how security is handled will to be provided.

What education is necessary for the personnel in external organisations, before they get access to information?

The safe haven personnel need to know the details of the agreement with PraksisNett and the GDPR requirements.

The researchers need to know how to delete records in research datasets that belongs to research subjects that have withdrawn their consents for participation and use of their information. The researchers also need to know the rights of the registered are handled by each of the joint controllers. They also need to know how to store data access and processing logs of research datasets to ensure that a person may request overview of who has used their personal information.

Third-party (i.e., Medrave software AB) personnel need to know the «The Code of Conduct for information security in the healthcare and care services". They also need to know the requirements stated in the agreement with UNN, which is responsible for processing.

Does the processor provide satisfactory guaranties for suitable technical and organisational measures that ensure that the processing will be performed in accordance with GDPR (Article 28 1)?

The data protection team at UNN will assess whether the specifications and the routines of the Snow team is acceptable and provide a recommendation to the GP practices.

For third party processors, performing operations on behalf of UNN, a specific agreement with specific requirements will be made. The third-party processors must satisfy these requirements.

Are all data flow, storage and temporal storage identified?

The data flow in PraksisNett is described in section 1.3.1.

As we have not completed the specification for safe storage of research datasets, all flow, storage is not identified.

How is the data transferred and made available (data flow)?

See section 1.3.1.

Where and how long is the information stored on the specific sources?

Local partial research datasets: The local research datasets stored on the Snow appliance box will be deleted when the data collection phase is completed and the partial research dataset is stored in the safe haven. Only the access and processing logs necessary to ensure the rights of the registered are kept (the DRC link between personal ID number and project ID and statistical processing logs).

Research dataset at the safe haven: Dataset at the safe haven will be stored until they have fulfilled their purposes. This include the time needed to ensure that the published research papers can be validated for correctness.

How long is the information stored after completing the purpose of the processing, before the data is deleted?

One day should be sufficient. When the data has fulfilled its purpose, it will be deleted.

When will the information be deleted?

The next day.

Have you developed routines for deleting information?

At the local data reuse server: No, not yet.

At the safe haven: Too early for that.

Safe storage of research datasets: Not yet

Is information security handled satisfactory?

The privacy protection team at UNN and the Data protection authority (Datatilsynet) has assessed whether the information security is handled satisfactory in the IT infrastructure and given it its recommendation.

Are all active and planned measures suitable to ensure confidentiality, integrity and accessibility of information about persons?

This cannot be evaluated by us.

Are security standards, policies and norms for information security followed?

For our part; yes. For third party vendor compliance needs to be checked regularly to ensure that standards, policies and norms are followed.

2.5 Assessment of the necessity and proportionality

2.5.1 Principles for data protection / privacy

Legal basis

Is the processing lawful, fair and transparent (article 5.1 letter a, and article 6 and 9)?

What is the legal basis for the processing?

- **Consent, agreement /contract, legal obligation, vital interests, exercising public authority, legitimate interest?**
- **Does the legal basis include the purpose of the processing and transfer to externals?**

The Snow appliance box is, when installed in a GP practices, part of the controllers (GP practices) responsibility and can be used for all legal purposes that follows from Norwegian law, especially the medical record law §6, which states that the patients´ data can be used in relation to medical treatment, internal control, patient administration and quality assurance of the health service.

The patients´ data on the Snow appliance box can also used for disease surveillance purposes, which is legal processing with a basis in the disease prevention law (smittevernloven) §7-1 and §7-2 b and e.

For research purposes written consent from the data subjects will be collected for use of the health information and for study participation. Researchers may also apply for exempt from getting consent from the registered from the regional ethical committees. However, such exempt may not be accepted by the PraksisNett board as PraksisNett has stated that patient data is only used with consent from the patient. Solutions for electronic consent from the research subjects may also become available.

Have you evaluated and controlled that the purpose is valid and reasonable?

- **What are the expected benefits of the processing? For the organisation, the registered and the society?**

- **Is there a clear separation between personal information that is necessary for the agreement and what information that need to be based on consent?**

The long term expected benefit from the processing and research is improved medical treatment for the patients, the health service and our society. Also, participation in research projects on specific diseases may be assessed as a benefit by the patients. Some patient may claim it is unfair that they cannot participate in research projects, because their GP is not part of the PraksisNett research infrastructure.

Research projects will define the variables that will be used in the research project. Written or electronic consent for use of the information will be collected from the patients. Exempt from consent may be given by the regional ethical committees but may not be accepted by PraksisNett board. The patients will be provided with a copy of the information from their electronic patient record used by the GP office, to ensure that the information is correct and that the patient is informed about what data that will be used in the research project.

Assess how transparency is handled in the processing.

The patient may ask and get information about what purposes his or her data has been used for. If requested, the processing logs of each data reuse project (including research projects) can be provided. When data has been transferred to safe havens, access logs and processing logs may be provided. However, processing performed in safe havens is outside the responsibility of the PraksisNett IT infrastructure.

Limitation of purpose (Formålsbegrensning)

The purpose of processing should be lawful, fair and transparent (article 5.1 b). Check the following:

- **Is the purpose clearly defined?**

The purpose of processing datasets within the PraksisNett IT infrastructure must always be specified. Each research project will pursue its specific objective.

Is the purpose defined in accordance with the expectations of the registered?

We assume that patients expect that their medical data is used to provide high quality medical services. The patients also assume that the information is used based on their consent.

Can the purpose be accomplished using less invading methods?

All processing in the Snow appliance box is based on pseudonymised and encrypted data. It is not possible to use anonymous data, as this would prevent ensuring that data subjects can exercise their rights (the right to withdraw consent, transparency of use).

Each research project will need to assess whether their purpose can be achieved by doing privacy preserving statistical processing on the data stored in the Snow appliance box.

Can the purpose be accomplished using anonymous or pseudonymous alternatives?

Each research project will need to assess whether their purpose can be achieved by doing privacy preserving statistical processing on the data stored in the Snow appliance box.

Data minimisation (article 5.1 c)

Personal data should be adequate, relevant and limited to what is necessary.

Assess the purpose of processing: Is it possible to achieve the purpose by:

- **Limiting the collection of personal information?**
- **Limiting the level of details of personal data**

- **Without using secret and sensitive information?**
- **Using aggregated or pseudonymous personal data?**

Assess the necessity and relevance related to the purpose for every variable in a dataset.

The variable set proposed by researchers will be assessed by the management board of PraksisNett as part of the project approval process. As part of this assessment, evaluation of these issues will be performed. Also, GPs will perform the same assessment when consenting to participation in a research project. However, the GP will potentially put a lot of trust on the management board's evaluation of these aspects.

Correctness (article 5.1 d)

Personal data should be correct and updated

How is personal data kept correct and updated, with or without involvement of the data subject?

Patient involvement: As part of the invitation letter to the patient, a copy of the clinical information used in the research project will be provided. Error discovered by the patient must be communicated to their GP, which should then correct the information in the research dataset or in the EHR (if legal and possible).

Without patient involvement: Automated data quality checks will be used to provide the GP with a tool to improve the correctness of the clinical data stored in the PraksisNett IT infrastructure.

Assess whether necessary functionality exist to correct or erase personal data.

Correcting a medical record may be difficult. We are uncertain whether corrections can be made in the EHR system. Legal advice is needed here. Error in clinical data need to be corrected by the GP responsible for the patient. Information should ideally (if possible) be corrected in the source, the EHR record. If that is not possible, data can be corrected in the data reuse server.

From the data subjects point of view, is there a need for contradiction?

Research subjects will be provided with a copy of the EHR information used in the research project. The research subject / patient may ask the GP to correct errors in the information.

Do the data subjects have the opportunity to object on what controller has registered?

Yes, by using the information provided as attachment to the study invitation letter. For information collected by the researcher or others, the researcher is responsible for handling this issue.

Do you have routines for how employees write journals, memos, minutes etc.?

No, but routines for writing medical record notes may need to have some.

Limitation of storage (article 5.1 e)

Personal data should be deleted or anonymised when the purpose has been accomplished.

Is the personal data kept after the purpose has been accomplished?

When is the data deleted?

When is the personal data pseudonymised or anonymised?

Research datasets that has fulfilled its purposes will be deleted. However, we assume most datasets will be stored in safe havens, outside the responsibility of the PraksisNett. Agreements between PraksisNett and the safe haven organisation will ensure that the responsibility to delete research datasets is placed with the researchers and the safe haven institution.

All data stored on the Snow Appliance box in the PraksisNett IT infrastructure will be pseudonymised.

Safeguards (Article 89.1)

What safeguards must be in place if personal data should be stored for a longer time period for purposes related to archives in public interests, purposes for scientific or historical research or statistics?

Before data can be stored in a safe haven, an agreement needs to be made between the researcher and the Safe haven. The agreement should state what policies will be used to safeguard personal data and how long personal data will be stored.

The research datasets may need to be stored for some time to ensure that published research can be validated for correctness. However, when research datasets have been deleted in the safe haven environment, a notification should be sent to the GP practices, notifying the GP practices that the specific research dataset has been deleted. If the data subject asks for an overview of what their data has been used for, information can be provided that the specific research dataset has been deleted. As long as such notification has not been sent, the data subjects can get information about where their clinical information is stored from their GP or the researcher.

Is it necessary or possible to improve the proposed methods of processing?

We don’t think this is possible today. It will always be necessary to improve the processing methods to reduce risk of exposing sensitive data.

The development of distributed secure multiparty computation methodology will at some time in the future remove the necessity of moving research datasets out of health institutions. Datasets can be analysed locally, while ensuring the privacy of patients and health personnel and minimising distribution of highly sensitive information.

Is it possible to improve the condition for the registered?

Only by advancing the distributed secure multiparty computations, as explained above.

2.6 The rights of the data subjects

Assess how the rights of the data subjects is handled:

Assess how information is provided to the data subjects? (article 12,13,14)

The patient will be informed by their GP practice and the researcher about their rights. They will also be informed about how and for what purposes their EHR data is used, who is able to access it, while stored in the EHR system, on the PraksisNett IT infrastructure, and in safe havens. We assess these efforts to be adequate and reassuring to maintain the trust between the patient and the GP and between the patient and the researcher, necessary to get consent for use of the information for research purposes.

When their data is used for research, additional information about how their data is handled, who will access it and where their data will be transferred will be provided to the patient as attachments to the invitation letter and consent forms.

Assess how consent is collected (Article 7 and 8)

Is the consent provided freely and explicitly?

Is the consent documented?

Consent for participation in research and use of personal data will be mailed to the patient. The patient must then sign and return the written consent to the researcher or to the GP practice. The researcher or the GP practice will store the consent forms. In the future electronic consent may also become possible.

Can the data subjects withdraw consent, as easy as giving it?

Yes, patient can withdraw their consent at any time by contacting the researcher or their GP. Which one depends on the design of the study and who invited the patient to participate in the study. The patient will be informed about who to contact in the study invitation letter.

Can the data subjects get insight into personal data?

Yes, the GP will be able to provide the patient with a copy of the data from the electronic patient record as part of the study invitation letter or later while the research dataset is still stored at the GP practice Snow appliance box. The researcher will be responsible for ensuring the same right towards the patient, if additional data is collected by the researcher.

When the research dataset is moved to the safe haven, dependent of the design of the study, the GP or the researcher will be responsible and must notify the safe haven and the responsible researcher for the research project to provide a copy of the pseudonymised data.

Can the data subjects get a machine readable copy of the data? (article 15 and 20)

Yes, the PraksisNett IT infrastructure will contain functionality to extract machine readable data about a specific patient.

Can the data subjects get his data corrected? (Article 16)

Yes, as long as the data is stored at the Snow appliance box. After transfer of the research dataset to a safe haven, the GP must notify the responsible researcher that information in the dataset need to be corrected for the specific pseudonym belonging to the patient. However, this issue is the responsibility of the safe haven institution.

Can the data subjects get his data deleted? (Article 17)

Yes, by contacting the GP that invited the patient. While the research dataset is stored at the GP practice the GP himself can delete the data in the research dataset. When the research dataset is moved to the safe haven, the GP must notify the responsible researcher to delete the data corresponding to the pseudonym of the patient.

Can the data subjects right to restrictions of processing, objections to processing, and right to notification regarding rectification or erasure of personal data be handled? (Article 18, 19, 21)

The patient can reserve himself from being invited to participate in research projects by notifying his GP. The DRC will then remove this person from inclusion in recruitment lists used to recruit patients to research studies.

Are automated individual decision making, including profiling, performed? (Article 22)

Yes, in the sense that patients are selected to be part of research datasets based on inclusion and exclusion criteria specified by a researcher. The GP may also use his local PraksisNett client to identify the patient based on information stored in the EHR system.

What is the legal basis for such processing? Explicit consent or law?

Both legal and consent.

Is it necessary or possible to improve the way the data subject rights are handled?

During the study invitation phase invitation letters, consent forms and clinical information will be provided to the patients through mail or through electronic communication. To perform these manual operations, support staff at the GP practice may be used. An assessment must be made whether such personnel can take part in this process. We need advice on this questions from our evaluators.

2.6.1 The data subject's freedom

Assess how the data subject's rights in relation to the European human rights convention is assured?

The right to privacy and net neutrality

The right to privacy has been realised as far as current technology is able to bring us. Advancement on distributed secure multiparty computations can bring us even closer to an environment where the risk of privacy breaches for the patient can be minimised.

We do not see how net neutrality is affected or affect us.

The right not to be discriminated

Patients do not have the right to participate in research on conditions they suffer from. However, they have the right to select their GP and to some extent the health institution where they will be treated. An assessment must be made on whether patients having GPs that do not take part in research studies on the patient's conditions is a way of discriminating the patients wish to contribute to research on their medical condition.

Freedom of thought, belief and religion

The patient is able to deny participation in research projects that have contact points with his/her strong religious views. Like abortion, vaccines, sexual preferences.

Freedom of speech and information

We cannot see any relation to these issues.

Is it necessary or possible to improve the way the data subject freedom is handled?

It is possible to give more patients in Norway access to participate in research projects on their medical condition. It is also possible to improve protection of patient privacy and reduce the exposure of their medical record data. However, this is a question of funding research within these domains. We assess the efforts described in this document to be adequate and reassuring to maintain the trust between the patient and the GP, necessary to get consent for use of the information for research purposes.

2.7 Assessment of the risks to the rights and freedoms of data subjects

In the sections above, we have described the plans and actions made to ensure the rights and freedom of the registered. In the sections below we will describe the reasoning behind the choices made and how these choices minimize the risks, along with the potential risks introduced by not following them may have for the patients.

The Snow system, which is the basis for the PraksisNett IT infrastructure has been subject to 8 risk assessments processes both internally and externally. A summary of these risk assessments is provided as an attachment to this document. The latest risk assessment report is provided as attachments to this document.

2.7.1 Measures envisioned to ensure privacy for the patient

At the core of ensuring the rights of the registered is to know where data about patients is stored and what the data is and has been used for. By making copies of health data and distributing it, we increase the risk of violating the rights and freedom of the registered. We also increase the cost for our society and for the patients to ensure that the rights to control what data exists and what it is used for. The principle followed in the PraksisNett IT infrastructure is therefore to minimize the number of copies made of health data. Data about patients will therefore be stored only two places, in the health institution it was created in and in safe havens where researchers can perform advanced statistical analysis on the data.

To minimize the risk of violating the privacy of study participants we have implemented a pseudonymisation mechanism for all patient data in the IT infrastructure. The key to re-identify the study participant will only be stored on the EHR server, together with the medical record. By storing the key together with the medical record, we ensure that participation in a medical study can be kept private and part of the patient – medical doctor relationship. Nobody else needs to know that the patient is participating in a research project. The reason for doing it this way is that participation in a research project can be as sensitive as the content of the medical record. Fear of exposing participation in a research project on a patient´s condition is one obvious reason for rejecting participation. Storing the consent form outside the health institution that recruited the patient should therefore be avoided. It should ideally be kept as part of the medical record of the patient. It is also essential that the medical personnel know that the patient is or has been participant in a study.

2.7.2 New risks introduced by the PraksisNett research infrastructure IT solution

As identified in the latest risk assessment the following threat are evaluated to have moderate risk:

- R9: The data reuse server can be stolen by persons with physical access to the GP practice.
- R13: The data reuse server may be infected by malicious software like viruses, trojans etc. resulting in server breach.
- R19/R21: The data reuse server may be breached through software vulnerabilities (so called supply chain attacks)

As identified in the latest risk assessment the following threat are evaluated to have low risk:

- R2: Unauthorised access to EHR data.
- R3: Unauthorised access to centralised servers
- R7/R8/R10/R12: Unauthorised persons get access to the GP practice network through remote system management tools.

For a discussion of all 21 identified threats please consult the latest risk assessment report.

2.7.3 Measures envisioned to address new risks

In section 2.7.2, the re-identification risk from aggregated data is listed. It is an old risk, but appears in the new IT solutions. We plan to address this risk by applying functionality for disclosure control before visualising data to the end users or making it available through web service interfaces.

In section 2.7.2, several new risks to violation of privacy are identified. Because the first version of the functionality to authorise access to the re-identification interface of the DRC will be done by the PraksisNett staff, a possibility exists that unauthorised personnel get access to fully identified patient information belonging to a research dataset at the GP practice during the recruitment phase. Two measures can be used to avoid this risk, 1) not authorising GP practice staff for such access, 2) establishing quality assurance routines that ensure that access is granted to personnel that is approved by the responsible GP.

Another possible risk related to quality of provided data is the risk of identifying wrong patients. The cause of such an event would probably be programming errors done by the PraksisNett development team. This event will be discovered very soon as GPs and GP practice staff will notice the mismatch between the recruitment criteria of a research project and the content of the recruitment list. If such a mismatch is discovered the PraksisNett team will be notified and the error corrected.

2.7.4 Measures envisioned to ensure the right not to be discriminated

As mentioned in section 2.6.1, patients do not have the right to participate in research that can improve the treatment of future patients. Some patients view participation in research on their medical condition as a deeply meaningful and valuable action. For some patients the only hope for finding a cure or treatment is participation in research. Every GP should therefore make an assessment whether not being part of the PraksisNett research infrastructure is discriminating the patient on his list from contributing to establishing new knowledge about their medical condition.

3 References

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). 32016R0679, http://data.europa.eu/eli/reg/2016/679/oj/eng (2016, accessed 29 June 2018).

[2] Vurdering av personvernkonsekvenser (DPIA). *Datatilsynet*, https://www.datatilsynet.no/regelverk-og-skjema/veiledere/vurdering-av-personvernkonsekvenser/ (accessed 13 June 2018).

[3] Bellika JG, Henriksen TD, Hurley JS, et al. Electronic health record data reuse infrastructure requirements - Ehealthresearch.no, https://ehealthresearch.no/prosjektrapporter/electronic-health-record-data-reuse-infrastructure-requirements (accessed 5 October 2017).

[4] Bellika JG, Henriksen T, Yigzaw KY. The Snow System – A Decentralized Medical Data Processing System. In: *Data Mining in Clinical Medicine*. Spinger, 2014.

[5] Hailemichael MA. Emnet: A System for Privacy-preserving Statistical Computation on Distributed Health Data, http://www.ub.uit.no/munin/handle/10037/9154 (2015, accessed 31 August 2016).

[6] Hailemichael MA, Marco-Ruiz L, Bellika JG. Privacy-preserving Statistical Query and Processing on Distributed OpenEHR Data. *Stud Health Technol Inform* 2015; 210: 766–770.

[7] [The Helsinki Declaration of the World Medical Association (WMA). Ethical principles

of medical research involving human subjects]. *Pol Merkur Lek Organ Pol Tow Lek* 2014; 36: 298–301.

[8] SAFE = bedre personvern i forskningsprosjekter. *Universitetet i Bergen*, https://www.uib.no/nb/foransatte/96464/safe-bedre-personvern-i-forskningsprosjekter (accessed 28 June 2018).

[9] University of Oslo. An introduction to TSD - University of Oslo, https://www.uio.no/english/services/it/research/sensitive-data/about/introduction.html (accessed 10 October 2018).