DANISH NATIONAL GENOME CENTER

# Minutes

## 3rd Meeting in the DNGC's international advisory board

Date: 30-05-2023
Unit: NGC
Caseworker: CSKA.NGC
Ref.: 2300208
Doc. No.: 2714700

**Date: May 11, 2023, 10.00-16.00 (CET)**
**Location:** Danish National Genome Center, Copenhagen

**Agenda**

| Item | App. time | Activity |
|---|---|---|
| 1/9 | 10.00 | **Welcome and presentation of the agenda**<br>/Tim Hubbard |
| 2/9 | 10.10 | **Status Danish National Genome Center**<br>/Bettina Lundgren |
| 3/9 | 10.45 | **Inputs on the Danish work with a renewed strategy**<br>/Lene Cividanes |
| 4/9 | 11.30 | **DNGC in the overall Health Data Landscape**<br>/Tim Hubbard |
| 5/9 | 13.00 | **Building and implementing a variant database – Strategies for reprocessing**<br>/Majbrit Hansen, Stein Karlsen and Nicolas Rapin |
| 6/9 | 13.45 | **Considerations on platform and cloud solutions**<br>/Ali Syed, Nicolas Rapin and Camilla Borchorst |
| 7/9 | 14.35 | **Cost benefit and clinical impact of WGS – WGS vs other methods**<br>/Birgitte Nybo Jensen, Malene Rasmussen and Peter Johansen |
| 8/9 | 15.30 | **Experiences on and suggestions for public-private partnerships**<br>/Lene Cividanes |
| 9/9 | 16.00 | **Next steps and wrap up**<br>/Tim Hubbard and Bettina Lundgren |

**Participants**
Tim Hubbard, Professor, Kings College London (Chair)
Richard Rosenquist Brandell, Professor, Karolinska Institute (Vice Chair)
Valtteri Wirta, Dr., Ph.d., Karolinska Institute
Dag Erik Undlien, Professor, M.D., PhD, Oslo University Hospital
Aarno Palotie, M.D., Ph.d, Institute for Molecular Medicine Finland
Ruben Kok, Ph.d., Director, Dutch Techcentre for Life Science
Kym Boycott, Clinical geneticist, University of Ottawa (online)
Jean- François Deleuze, Ph.d., Head of CNRGH (online)
Heidi Rehm, Ph.D, Broad Insitute (online)

**DNGC's Secretariat**
Bettina Lundgren, Director, DNGC
Christian Dubois, Chief of Staff, Management Secretariat, DNGC
Lene Cividanes, Head of Research and International Relations, DNGC

Ivana Bogicevic, Policy officer, Research and International Relations, DNGC
Christiane Karle (secretary), Academic officer, Research and International Relations, DNGC

**Minutes**

**1/9 Welcome and presentation of the agenda /Tim Hubbard (Chair)**

Tim Hubbard welcomed everyone and introduced the program shortly.

**2/9 Status Danish National Genome Center /Bettina Lundgren**

Bettina presented a short status of DNGC. Currently, the National Database contains 13,413 genome samples. The samples are collected from all 17 patient groups. DNGC is in the process of setting up a framework for evaluating the effect of whole genome sequencing (WGS) in the patient groups.

After Bettina's presentation, the following was discussed:

- DNGC was advised to aim at **retrieving relevant health data or facilitate access to relevant health data on the individual patient**. The new strategy must address how a centralized data system can be organized facing two ways: clinical and research.
- The system should be organized and based on **pseudonymization by using personal "key identifiers"** that links data and multiple environments.
    - These two points are currently also key considerations in the "Danish vision for better use of health data". Research projects are to be connected, so that clinical data with anonymization as well as meta data can be reported to researchers for analyses through flagged cases/ID numbers, and vice versa.
    - As part of the "Danish vision for better use of health data" DNGC and Denmark are most likely heading towards a **hub-and-spoke model/a federated system** with data from relevant agencies. It is currently heavily discussed in Denmark; which authorities should be involved and also who should be responsible - and if legislation needs to be modified.

**3/9 Inputs on the Danish work with a renewed strategy for personalized medicine /Lene Cividanes**

The Danish government is planning to present a new national strategy for personalized medicine in 2024. The strategy will build upon and be a continuation of the former strategies (2017-2020 and 2021-2022). Outcome of the first discussions on key elements in the strategy conducted in the "Steering group for Personalized Medicine" was: Effective treatment trajectories for patients; Better usage of health data; and Strengthening the competencies of healthcare professionals.

The presentation was followed by a joint discussion in order to obtain inputs from the advisory board on relevant elements to include in the coming strategy.
The following points were discussed:

- **Development of a clear framework for Evaluation of health economics/economic impact of WGS** is key for Denmark and for the regions when planning the future implementation of WGS in health care and when taking decisions on which genetic analyses to be used in which disease areas in the future.
- DNGC could consider **implementing a structure/directory for helping the regions monitor the implementation of WGS at the patient group level**. This can be done by using Danish registries, because it is possible to extract and monitor e.g. how WGS in certain disease areas have caused documented fewer visits to health care (i.e. show savings in a formal way) etc.
- To ensure public legitimacy and understanding of genetic analyses **engaging patients** in deciding on how, to whom and to which level the WGS offer in the Danish Health care setting should continue is key. It could be a e.g. in the form of a patient and/or citizen panel. It was further discussed if patients benefitting from a public health care system can be expected to contribute to other patients' treatment (to a certain extent).
- Very advisable to include more focus on **pharmacogenetics** and **prevention** in the future– as much data as possible should be collected and made accessible to clinicians and researchers at the DNGC.
- The new strategy must be formulated in parallel to and **taking into consideration EHDS.** As EHDS aims at creating a unified structure, involved countries need to have a federated system ready. DNGC has to think both in short term and long-term strategies. If countries start building federated systems now it will also "put pressure" on EHDS and make things happen faster.
- Currently, data can in principle only be accessed from a Danish IP address by Danish based researchers – this may limit Denmark in being part of global landscape. DNGC **need to consider to open up for access from other countries** (EHDS will also require this), especially in terms of interpretation and sharing information on variants.
- **Genome of Europe/a National reference cohort**: The new strategy should aim at including the establishment of a **research cohort** or **a national reference cohort** (rather than a reference genome or a GRAF genome). There are already a number of existing research cohorts/longitudinal studies for complex areas/diseases in Denmark: These cohorts could be brought together and linked to genomic data.
- An **"obstacle removal process"** was suggested (a list of necessary solutions to overcome obstacles). This could help engaging the ministry to build cases and policies on "the obstacles" – and structure the solutions towards what society wants in terms of data use and within the legal possibilities (or adapted legal possibilities, e.g. new laws may be needed to overcome obstacles). The infrastructure should be built or adjusted according to these policies.
- Since DNGC was established, main focuses have been on implementation of WGS in health care and establishing a national framework (legislation, governance, education etc.) and a national IT-infrastructure, bridging in the regions and with the national level. The new strategy must focus on **interpretation, utilization and optimization, i.e. data integration, research,**

**implementation science, accumulation of knowledge and interpretation – show how collecting WGS data gives critical utility and value for money.**

- Suggestion on focusing on three main pillars in the strategy: 1) **implementation phase** – looking at tax payer funding, equal access to tests, enable treatment and patient benefits, ensuring cost effectiveness, 2) continued **support to pilot projects**, e.g. creating new cohorts, new patient groups, and involving new technologies – and include follow up strategies for cohorts, 3) **dimension for discovery,** i.e. data must be used for research, infrastructure and federation, data usability, linking with other data ("lifetime personalized medicine").
- Continued focus on **engaging the clinical genetic community – facilitate knowledge sharing, workshops across the country**
- As a national agency with a centralized genome database, DNGC must emphasize that size of data matters, as well as **ensure equal access to and continuously usage of data across regions**. The database should set up a system for continuous usage and care of data along patients' life path.

**4/9 DNGC in the overall Health Data Landscape /Tim Hubbard**

Tim Hubbard presented how DNGC is operating as part of the overall health data landscape (genomes, health data, non-health data, and the Danish health system). On a global level, DNGC is part of 1+ Million Genomes as well as the Genomic Data Infrastructure (GDI) and Denmark is also part of the pilot projects preparing for the European Health Data Space (EHDS). Countries participating in EHDS and 1+Million Genomes need to work towards federation, metadata catalogues and data standards in order to engage with each other.

Afterwards it was discussed how the international initiatives (GDI, B1MG, 1+MG) interact and differ, and how member countries need to prepare federated data systems. The OMOP Common Data Model was mentioned as an example for translation of data, e.g. for cohort discovery across different datasets.

**5/9 Building and implementing a genome- and variant database – Strategies for reprocessing /Majbrit Hansen, Nicolas Rapin and Stein Karlsen**

Majbrit Hansen and Nicolas Rapin presented how DNGC is in the process of building a genome and variant database. The genome database is getting its content from the General Purpose Reporting flow (GPR), which gathers data from genetic analyses (e.g. WGS) on patient samples and consists of metadata, and pipelines run information reported by the regions. Currently, DNGC receives samples from two sequencing facilities, in Copenhagen and Aarhus. The Genome Database will be developed further in the coming years in order to establish an infrastructure for storing samples and processed data, as well as combining data from other Danish registers and cohorts.

Afterwards it was discussed, what is the best strategy for storing data in terms of file formats. Currently, DNGC stores FASTQ files, which is a space-consuming solution. Other smaller file formats (e.g. CRAM and VCF files) may be more appropriate (cheaper) for long-term storage, while still keeping necessary information for reanalysis. In England, CRAM files are removed from active storage after a year, unless

clinicians are still analyzing patient data. In Sweden, FASTQ files are stored in a compressed format (SPRING) and then decompressed if needed.

In general, DNGC must not only focus on long term storage of raw genome data on individual level, but also focus on building robust pipelines and variant interpretation across compressed data, in "preprocessed" databases, as many future inquiries will be based on this.

**6/9 Considerations on platform and cloud solutions /Ali Syed, Nicolas Rapin, Camilla Borchorst**

Ali Syed presented the design and implementation of DNGC's infrastructure. DNGC operates the second-largest HPC infrastructure in Denmark with an automated secure cloud service environment and providing reference tools and datasets. DNGC has started to collaborate with other Danish health agencies (Danish Medicines Agency and Statens Serum Institut) on collecting different types of data profiles and pipelines in one system.

Nicolas Rapin presented how the project TRANSLATE processes WGS's from diabetes patients in collaboration with DNGC using a modified Sarek nextflow pipeline. TRANSLATE is the first large-scale implementation of systematic genetic testing within a common, non-communicable, chronic disease in Denmark.

Afterwards it was discussed how DNGC and the regions have divided the data flow process and storage. Currently, the dataflow is divided in two layers: 1) DNGC as the data owner (only FASTQ files), and 2) the regions as the production responsible, responsible for the pipeline and storing data in several formats.

It was suggested, that DNGC impose restrictions on the regions' storage and in collaboration with the regions write a data management plan that specifies plans for long term storage (e.g. VCF files are kept forever, CRAM file for a certain period, etc.). Furthermore, it was recommended that log files are kept forever.

**7/9 Cost benefit and clinical impact of WGS – WGS vs other methods /Birgitte Nybo Jensen, Malene Rasmussen and Peter Johansen**

Malene Rasmussen presented DNGC's model for evaluating the implementation and effect of whole genome sequencing among the 17 selected patient groups. The evaluation is, among others, based on a literature review, patient cases, semi-structured interviews with clinicians and patient representatives, and international experiences. The evaluation does not include data on clinical impact or diagnostic yield, since DNGC as a government agency does not have direct access to clinical information in the regions.

Afterwards, it was discussed and emphasized by the advisory board that one of the most important challenges related to the evaluation of the value of WGS in healthcare is that DNGC does not have access to clinical data such as diagnoses or health outcomes. Without clinical information, the effect of WGS in healthcare cannot be evaluated sufficiently. The advisory board therefore suggested that it should not be called an *evaluation* – rather a report on i.e. "initial experiences". In order to allocate resources going forward, DNGC should in collaboration with the ministry

and representatives from the regions assess how effects of genomic analysis including whole genome sequencing can be measured and evaluated, also in the long term. Furthermore, an interregional agreement should be prioritized in order to share genomic and clinical data across regional borders.

**8/9 Experiences and suggestions for public-private partnerships /Lene Cividanes**

DNGC is currently looking into how to put together a realistic model for engaging in public-private partnerships being a public agency in a Ministry working with genomes. There are several prerequisites related to potential partnerships; Security, legislation, national policies/political interests etc. Currently, DNGC has not engaged in public-private partnership but only in public funded projects.

A public-private partnership could be centralized around how to develop new technology, innovation or research for better prevention, diagnostics, and treatment opportunities.

It was discussed how DNGC can identify areas of collaboration and initiate public-private partnerships. DNGC was advised to be careful with initiating partnerships within core services and not rely on one supplier, but rather prioritize a few, smaller partnerships.

DNGC needs to identify the investment window and critical needs for technology and innovation to provide best service to patients. It was also mentioned that DNGC should consider how the Danish pharmaceutical industry can be involved in order to co-develop areas within life science and precision medicine.

**9/9 Concluding remarks /Tim Hubbard and Bettina Lundgren**

Tim Hubbard and Bettina Lundgren closed the meeting and thanked everyone for the many good inputs from the sessions.