NORDIC CLINICAL GENOMICS DATA SHARING WORKSHOP 2
21.-22.11.2016

Workshop summary report

Oslo Universitetssykehus
SciLifeLab
Karolinska Universitetssjukhuset
University Hospital Copenhagen (Rigshospitalet)
DNV GL

Report No.: 2, Rev. 2
Date: 2017.02.07

Project name:	Nordic Clinical Genomics Data Sharing workshop 2 21.-22.11.2016
Report title:	Workshop summary report
Participants:	Oslo Universitetssykehus SciLifeLab Karolinska Universitetssjukhuset University Hospital Copenhagen (Rigshospitalet) DNV GL
Date of issue:	2017.02.07
Revision:	2

Objective:

The objective of this report is to summarize topics discussed and agreed actions from the Nordic Clinical Genomics Data Sharing workshop at SciLifeLab's offices in Stockholm 21.-22. November 2016 between the above listed partners.

Prepared by:

Guro Meldre Pedersen
Principlal Researcher
DNV GL

Sharmini Alagaratnam
Principal Researcher
DNV GL

Rev. No.	Date	Reason for Issue
1	2016.12.05	Draft summary report
2	2017.02.07	Final summary report

Table of contents

1 SUMMARY OF AGREED ACTIONS 1

2 EXECUTIVE SUMMARY 1
2.1 About the workshop - background 1
2.2 Workshop participants 1
2.3 Goals of collaboration 1

3 UNDERSTANDING LEGAL BARRIERS AND ENGAGING WITH KEY STAKEHOLDERS 2

4 DEFINING DATA, TOOLS AND PROCEDURES TO SHARE 5
4.1 Discussion on different levels of variant sharing 5
4.2 Sharing of variant interpretation procedures 6
4.3 Sharing of Ella variant interpretation tool 6
4.4 Actions 7

5 EXISTING AND NEW VEHICLES FOR SHARING 8
5.1 Existing variant classification sharing solutions 8
5.2 Anonymized Nordic frequency database 9
5.3 Sharing of curated (single) variants with associated phenotypes 10
5.4 Actions 11

6 ENHANCE QUALITY OF DATA AND PROCESSES 13
6.1 Benchmarking and variant interpretation 13
6.2 Actions 20

7 NORDIC NETWORK FOR CLINICAL GENOMICS 21
7.1 Defining the Nordic added value 21
7.2 Organization of further collaboration – defining the group 22
7.3 Actions 22

8 NEXT WORKSHOP 23

APPENDIX 1: WORKSHOP AGENDA 24

APPENDIX 2: LIST OF PARTICIPANTS 26

APPENDIX 3: HARMONIZATION OF CLINICAL INTERPRETATION OF VARIANTS 27

APPENDIX 4: SLIDES USED DURING THE WORKSHOP 28

1 SUMMARY OF AGREED ACTIONS

During the workshop discussions, several follow-up actions were identified. This table summarizes the actions identified in the sessions described in further details in sections 3-8 of this workshop summary report.

Table 1 Summary of actions agreed during the workshop

Theme	Action	Responsible	Deadline
Benchmarking	Run analysis and share results	Rigshospitalet/ Wiktor	01.03.2017
Benchmarking	Share annotated outputs	OUS/ Svein Tore SciLifeLab/ Måns Rigshospitalet/ Wiktor, Ane	01.03.2017
Benchmarking	Compare variant calling, annotation results	DNV GL/ Sharm	01.03.2017
Benchmarking	Run gene panel-specific calling	DNV GL/ Sharm coordinate	01.03.2017
Benchmarking	Run and compare ranking results: true case?	DNV GL/ Sharm coordinate	Backburner 01.03.2017
Benchmarking of variant interpretation	Benchmarking of variant interpretation per ACMG guidelines between OUS and Rigshospitalet	OUS/ Morten Rigshospitalet/	01.04.2017
ClinVar	Submit variants to ClinVar	OUS/ Svein Tore SciLifeLab/ Måns Rigshospitalet/ Wiktor, Ane	01.03.2017
ClinVar	Next WS: Comparison on experiences	All labs	April 2017
Collaboration	Identify relevant Finnish institution / department and consider inclusion in collaboration	SciLifeLab/ Valtteri	01.04.2017
Collaboration	WS I summary report to be distributed to new members of network	DNV GL/ Guro	22.11.2016
Collaboration	Distribute Norwegian Strategy for Personalized Medicine to members (SE, DK)	DNV GL/ Guro	31.12.2016
Collaboration	Name contest, winner to be awarded next workshop	DNV GL/ Guro to coordinate	18.04.2017
Collaboration	Propose principles for cooperation - Governance statues - Rules of collaboration incl. policy statements reflecting agreed goals - Confidentialities Openness	DNV GL/ Vibeke	18.04.2017
Collaboration	Value statement Nordic cooperation	DNV GL/ Guro	18.04.2017
Comparison of population-specific variants	Eidi to share 38 non-ExAC variants with Henrik for him to check if also present in Swedish databases	OUS/ Eidi SciLifeLab/ Henrik	01.03.2017

Theme	Action	Responsible	Deadline
Comparison of population-specific variants	Query SweFreq with own variants	OUS/ Svein Tore SciLifeLab/ Måns Rigshospitalet/ Wiktor, Ane	01.03.2017
Data governance	Invitation to pilot DNV GL Data quality assessment framework;	DNV GL / Stephen	01.03.2017
Harmonization of clinical interpretation of variants	Summary of variant interpretation pipeline for sharing	Karolinska/ Nicole	01.03.2017
Harmonization of clinical interpretation of variants	Sharing of SciLifeLab ranking process	SciLifeLab/ Måns/Henrik	Next workshop (agenda item) 01.03.2017
Legal	All workshop participants invited to contribute to establishing a Nordic network for legal competence on sharing of genomic data by identifying and sharing resources on the subject (relevant for BigMed WP5).	All	Open
Legal	Approach Norwegian Health Directorate for clarifications on Nordic variant frequency database	DNV GL/ Guro	12.12.2016
Legal	Identify regulating authorities in NO, SE, DK from which to seek approval for sharing of variants	SciLifeLab/ Valtteri OUS/ Dag Rigshospitalet/ Morten	01.04.2017
Next workshop	Identify venue in Copenhagen	DNV GL/ Guro	31.12.2016
Next workshop	Gather input and set optimal dates after Easter 2017	DNV GL/ Guro	12.12.2016
Next workshop	Gather input & develop draft agenda	DNV GL/ Guro	04.04.2017
Next workshop	Each group to define participants		
Nordic variant frequency database	SweFreq database and tools sharing	SciLifeLab/ Henrik	01.04.2017
Nordic variant frequency database	Identify and access test dataset for aggregation	All labs OUS: SciLifeLab: Rigshospitalet:	01.04.2017
Nordic variant frequency database	Test SweFreq aggregation tool, feedback to developers	All labs OUS: SciLifeLab: NA Rigshospitalet:	01.04.2017
Platform	Nordic cooperation forum on slack.com	Rigshospitalet/ Wiktor	01.03.2017
Sharing of variants – CASE – all variants (VCFs) and phenotype	CASE-development; 1-pager defining scope and approach for sharing VCFs and phenotypes: access to Norvariome from other diagnostic labs, legalities. Obtain provision for defined group. To be used as basis for technical and legal discussions.	To be discussed further	

Theme	Action	Responsible	Deadline
Sharing of variants – CASE - Curated single variants + phenotype	CASE-development; 1-pager on sharing of curated (single) variants, also to be used as basis for legal discussions	To be discussed further	
Sharing of variants - CASE - Population variant frequencies	Nordic variant frequency database CASE-development; 1-page document defining scope and approach for Nordic variant frequency database, to be used also for legal clarifications	DNV GL/ Sharm	21.02.2017
Tools – ELLA	Sharing of ELLA code	OUS/ Svein Tore	25.11.2016
Tools – ELLA	Testing of ELLA	SciLifeLab/ Måns Rigshospitalet/ Wiktor, Ane	01.04.2017
Tools – ELLA	Definition of format for communication between Ella and Scout	OUS/ Svein Tore SciLifeLab/ Henrik	01.04.2017

2 EXECUTIVE SUMMARY

This report summarizes the workshop that was held in Stockholm 21.-22. November 2016 between the below parties, focusing on sharing of clinical genomics experiences, tools, procedures and data. Supporting slides used during the workshop are available in appendix 4 of this document.

2.1 About the workshop - background

This workshop was a follow-up to the initial workshop taking place between the parties in Oslo 30.-31. May 2016, summarized in the report "Clinical Genomics Data Sharing – Workshop summary report".

2.2 Workshop participants

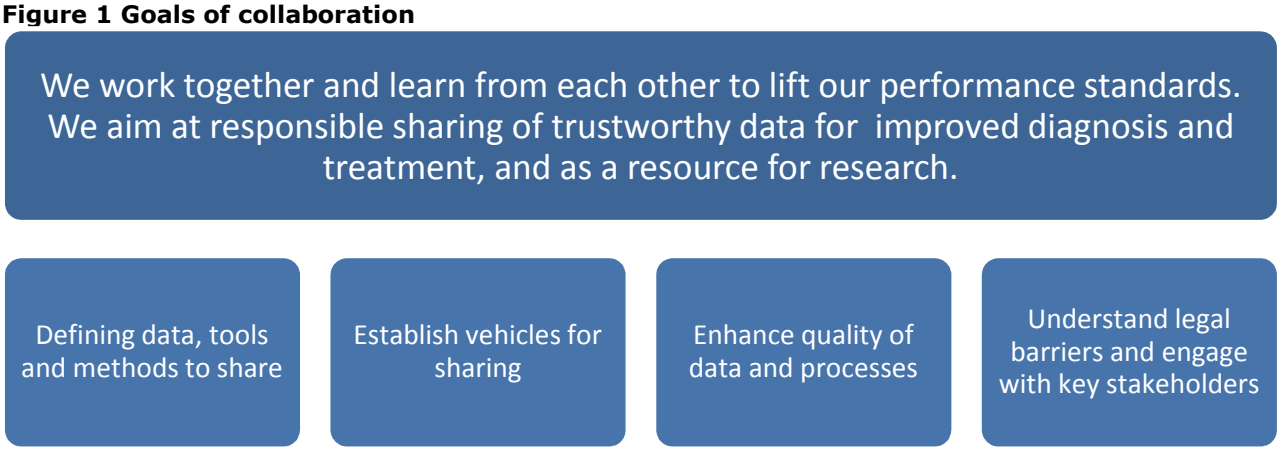
The workshop included representatives from the below units. A full list of participants is provided in Appendix 2.

- Department of Medical Genetics, Oslo University Hospital, Oslo, Norway.
- Clinical Genomics Unit, SciLifeLab, Karolinska Institutet, Stockholm, Sweden.
- Center for inherited metabolic disorders (CMMS), Karolinska University Hospital, Stockholm, Sweden
- Department of Clinical Genetics at The Juliane Marie Centre, Copenhagen University Hospital and the University of Copenhagen, Denmark.
- Center for Genomic Medicine at Diagnostic Centre, Copenhagen University Hospital and the University of Copenhagen, Denmark[1]
- DNV GL

2.3 Goals of collaboration

During the workshop the parties agreed to the below overall goals for the collaboration.

Figure 1 Goals of collaboration



[1] New to the collaboration; did not take part at the initial workshop in May 2016.

3 UNDERSTANDING LEGAL BARRIERS AND ENGAGING WITH KEY STAKEHOLDERS

Recognizing that ongoing initiatives and existing regulatory framework has a bearing on opportunities for sharing of clinical genomic data, the workshop included orientations on national initiatives, regulatory framework and relevant projects.

Table 2 Understanding legal barriers and engaging with key stakeholders

Agenda item		Main content	Discussion points

Nordic national strategies – status and goals relevant for the Nordic collaboration	Norway	Norwegian Strategy for Personalised Medicine in Healthcare[2] - developed by the Directorate for Healthcare per assignment from the Ministry of Health and Care services. The Directorate responsible for coordination of implementation.
		- Process timeline
		- Objectives
		- Strategic Areas
		  - Expertise and information
		  - Quality and academic and clinical development
		  - Health registries
		  - ICT
		  - Research & innovation
		- Main recommendations
		- Next steps
			- Strategy for 2017-2021 within current framework (funding, legislation)
			- Strategy focus on national / Nordic / international collaboration
			- BigMed specifically mentioned as one of the relevant projects supporting the strategy

	Sweden	- No national strategy under development in Sweden - Key players - SciLifeLab focus areas & collaborations - Major focus on formalising the collaboration between SciLifeLab and the Karolinska University Hospital to provide legal basis for further work. - International / European networking to identify potential partners for collaboration - SweFreq[3] published	- Bottom-up process for national focus on genomics - Identification of key stakeholders and processes for development of national strategy in the other Nordic countries - Finland is working on the implementation of the national strategy for precision medicine. The implementation is to be led by departments Ministries, and will focus on extensive sharing of genotype / phenotype data

[2] https://helsedirektoratet.no/publikasjoner/strategi-for-persontilpasset-medisin-i-helsetjenesten

[3] https://swefreq.nbis.se/#/

Agenda item	Main content	Discussion points
Denmark	- Next Generation Sequencing landscape in Denmark; major players - Copenhagen area a geographical hub for relevant stakeholders - The Genome Denmark[4] platform for sequencing and bioinformatics, a consortium of Universities, Hospitals and Industry partners - Strategy paper; Action plan for Precision Medicine[5] - Danish e-infrastructure in planning, decisions and funding pending.	- No political decision taken on implementation of precision medicine; ongoing discussions - Strategy paper[5] developed by the research side, not from the clinical perspective
Regulatory framework	EU General data protection directive / regulation - Timeline & objectives - National interpretations	BigMed WP 5 will focus on legal issues with the ojectives to: - Ensure the project is aligned with the emerging EU developments addressing protection of personal data and provide combined technical and legal input to national and international regulations under development. - Engage interested parties to establish a legal and ethical team that will act as a "centre of excellence" to support the project to operate within acceptable legal and ethical boundaries. - Ensure external and internal visibility of the legal and ethical dimensions of the project.
Relevant projects	BIG data MEDical solution – BigMed - Project financed by the Norwegian Research Council - Project period: 2016-2019 - Partners: Universities (technical and legal), hospitals, industry, patients. OUS, Karolinska University Hospital, SciLifeLab and DNV GL partners. - Vision: Lay the foundation for an ICT platform that addresses the analytical bottlenecks for the implementation of precision medicine and paves the way for novel big data analytics. The solution will provide the patients with an optimized care which takes their unique individual characteristics into proper consideration.	BigMed work packages relevant for the Nordic collaboration: - WP 0: BigMed community and solution concept - WP 2: Bioinformatics pipeline and molecular pathology - WP 3: Genomics and data sharing - WP 5: Legal and ethical considerations

[4] http://www.genomedenmark.dk/

[5] http://www.regioner.dk/media/1280/handlingsplan-for-personlig-medicin.pdf

Table 3 Agreed actions

Theme	Action	Responsible	Deadline
Collaboration	Identify relevant Finnish institution / department and consider inclusion in collaboration	SciLifeLab/ Valtteri	01.04.2017
Legal	All workshop participants invited to contribute to establishing a Nordic network for legal competence on sharing of genomic data by identifying and sharing resources on the subject (relevant for BigMed WP5).	All	Open

4 DEFINING DATA, TOOLS AND PROCEDURES TO SHARE

4.1 Discussion on different levels of variant sharing

In this session, participants were divided into four groups corresponding to four levels of genomic data that could potentially be shared. Each group was to identify for each data level: grounds for sharing, existing means for sharing, and what added value Nordic sharing would generate. After discussion, it was agreed that the lowest level of data to share would be modified to *curated/classified* single variants.

A catalogue of variants in hypernormal controls (specifically individuals with no known pathogenesis) had previously been named in Workshop I as a useful resource. This item is addressed by the frequency database (Section 5.2).

Table 4 Summary of discussion

Data type	Why share?	Existing solutions	Nordic added value	How to address
A. Population variant frequencies	- Good starting point for sharing, achievable - Improved frequencies of very rare variants - Potential for later expansion	- SweFreq,NorGene, Norvariome, Danish ref - International (research) databases - Internal databases	- Population specificity, even if non-nordic patients included - Insight not available from other sources	See Section 5.2
B. Curated single variants + phenotype	- Quickly identify disease-causing variants - Rule out variants as non-pathogenic - Aid in identifying misclassified variants	- Clinvar - Matchmaker exchange - Ella classification support tool	- Achievable	See Section 5.3
C. All variants (VCFs) + phenotype	- Creates larger contextual db - Improve diagnostics, best to trace back - Possible to discover new genes/modifiers	- Matchmaker exchange - Norvariome	- Specific to Nordic population - Trust in quality of data - Easy access to lab of origin - Concrete cooperation, knowledge dissemination	Develop as a case as basis for discussions on technical and legal issues. Norvariome as a test case to allow other diagnostic labs access? Provision for a defined group?
D. Full genome + phenotype	- GWAS-like studies req >> data - Find patients with same phenotype - Create reference genomes (hard!)	- Matchmaker exchange		BigMed to chart this area: PARKED for now

4.2 Sharing of variant interpretation procedures

See Table 13, Section 6.1 for detailed description of variant interpretation procedures and discussion on harmonization of variant interpretation methods, but briefly the following participants presented their procedures.

Table 5 Variant interpretation procedures shared

Institution	Participant	Presentation
OUS - Department of Medical Genetics	Morten Eike	20161121 OUS Clinical variant interpretation and ELLA.pdf
Rigshospitalet - Department of Clinical Genetics	Morten Dunø	20161121 Rigshospitalet Dep of clinical genetics - Interpretation.pdf
Rigshospitalet - Center for Genomic Medicine	Ane Yde Schmidt	20161121 Rigshospitalet Center for genomic medicine - interpretation.pdf
SciLifeLab - Clinical Genomics Unit	Måns Magnussen, Robin Andeer	Live demonstration
Karolinska - CMMS	Nicole Lesko	Live demonstration, summary of procedure to come

4.3 Sharing of Ella variant interpretation tool

A breakout session with participants from all three sites looked at and discussed how Ella could be shared and co-developed.

Table 6 Summary of discussion

Agenda item	Main content	Discussion points
Demonstration of Ella (Svein Tore)		
Use case definition	- Scout to send a variant query which Ella then runs and sends a response when result is ready for Scout to fetch - Ideally submit class 4&5 variants from Scout to Ella - Benign variants not classified	- Output to be frozen in Scout for continuity, does not exclude future runs - Ensure compatilbility for ClinVar submission (Variant + class based on ACMG criteria) - Requires agreement on annotation format

4.4 Actions

Table 7 Agreed actions

Theme	Action	Responsible	Deadline
Sharing of variants – CASE – all variants (VCFs) and phenotype	CASE-development; 1-pager defining scope and approach for sharing VCFs and phenotypes: access to Norvariome from other diagnostic labs, legalities. Obtain provision for defined group. To be used as basis for technical and legal discussions.	To be discussed further	
Tools – ELLA	Sharing of ELLA code	OUS/ Svein Tore	25.11.2016
Tools – ELLA	Testing of ELLA	SciLifeLab/ Måns Rigshospitalet/ Wiktor, Ane	01.04.2017
Tools – ELLA	Definition of format for communication between Ella and Scout	OUS/ Svein Tore SciLifeLab/ Henrik	01.04.2017

5 EXISTING AND NEW VEHICLES FOR SHARING

Tony Håndstad (OUS) gave a brief overview of existing sharing initiatives, and discussed the extent to which these meet current needs (5.1). Two new vehicles for sharing were discussed, for an anonymized Nordic frequency database (5.2) and the sharing of curated (single) variants with associated phenotypes (5.3).

5.1 Existing variant classification sharing solutions

Variant classification and case-level databases aim to capture knowledge from past cases and published literature, with many existing public and private alternatives. In brief, no single standard for sharing exists, though existing open-source database solutions can be tweaked to suit purposes. In-house developed solutions at OUS & SciLifeLab are not designed for sharing. Existing sharing solutions consist either of centralized (for frequency, variant classification and case-level databases) or federated local (GA4GH, Beacon, Matchmaker exchange) databases. BigMed has as its main aim to provide a higher level, multi-source sharing solution. Some concerns were raised about the possibility that the combination of Beacon, Matchmaker Exchange and ClinVar covered all existing needs, and that focus should be on using globally recognized tools.

Table 8 Summary of discussion on existing sharing initiatives

Sharing initiatives	Discussion
ClinVar	- Biggest issue is reliability: can entries be trusted? - Requires submittion of publication or documentation of classification - Laborous to submit classifications - However resource is free and public - Likely to become the largest variant classification database (Nov 2016: 240150 records submitted, 149957 with assertion criteria) - ClinVar best alternative for sharing classification?
BEACON – GA4GH	- Beacon network can be queried to check if variants are known at other sites - Either manually or in an automated manner via API
Matchmaker Exchange - GA4GH	- Query to trusted partners: phenotype and genotype queries - Patient case based query
Plan for BigMed	- Make a report on functional and technical requirements of genomic databases - Exchange frequency data between Nordic countries

5.2 Anonymized Nordic frequency database

This session centred around the utility and specification of a potential anonymized Nordic variant frequency database. Two-phase solution proposed, where the initial phase aims to aggregate frequency summaries from the three sites in a quick and easy manner. A longer term goal would be to set up a more accurate database that would enable cohort calls to be made, and in future may incorporate phenotypic annotation/be publically available should the legal framework around this allow.

Table 9 Summary of discussion

Agenda item	Notes
Value of database for indiv. partners	Eidi (NO) – rule out variants that are not disease-causing, in particular for early onset diseases – late onset diseases more difficult Morten (DK) – rule out variants that are not disease-causing – use to filter Henrik (SE) – will NOT use to filter out, but to rank and prioritize
Discussion around specifications	ExAC not good alternative as need to submit raw and full genome data Documentation on how variant calling was performed important to trust quality Minimum quality criteria for submitting data? E.g. OUS list of 'bad genes' with frequent technical artifacts (tech specific). Differ btw labs? Blacklist artefacts, not genes! Rule out noise, but tech artefacts may in fact mask pathogenic variants Of value to compare and identify technical artifact vs population differences for filtering
Initial phase	Ideally each lab to aggregate frequency with a single tool (SweFreq?) Then share aggregated freq databases with other labs (check legal) Discuss aggregation of aggregated files DNV GL has possibility to host aggregated files (check legal) and develop functionalities Datasets: AMG in-house data (NO), SweFreq(SE), Danish trios (DK)
Specifications	**Must-haves** Quality data Technology used to produce data, versions Regions included (BED files) Tracking of included samples (avoid double inclusion) Site of origin (lab / database) **Should NOT include** Person identifiable data

Agenda item	Notes
Systems architecture	Aggregated vs federated databases, two very different needs and solutions: • Beacon – federated search on global variants to find the +1 case • Aggregated database: no need to update as frequently, used to rule out Update frequency: live vs periodic Security Access and authorization Possible solution • Each site create aggregated vcf: encrypted transfer or kept locally • Hosted at consortium site (via internet) or at DNV GL

5.3 Sharing of curated (single) variants with associated phenotypes

The sharing of curated variants was identified as a useful effort. Single curated variants represent the simplest level for sharing, with sharing of all variants/whole genomes was seen as a stretch goal (Dag). As a result the consortium partners agreed to develop a case for sharing of curated variants (see section 0). This could potentially be achieved by annotating variants in the population frequencies database, either with HPO terms, descriptive language or class ratings. An alternative solution could be to use Beacons to find variants and accompanying classification (Tony).

Table 10 Discussion - case: sharing of curated (single) variants & phenotypes

Main content	Discussion points
Variant interpretation information	Need for harmonization - OUS and Rigshospitalet base the interpretation on the ACMG guidelines - SciLifeLab / Karolinska use different approach
How to share – ClinVar & other databases	Challenges of current databases - Uncertainties in quality → must check content - A database of consistent high quality would be trustworthy and reduce work ClinVar - existing resource; "becoming too big to fail" - Based on ACMG guidelines Consensus on testing of ClinVar contributions - All labs committed to sending variants to ClinVar and compare experiences; does this cover the needs? - SciLifeLab will introduce automatic submitting to ClinVar - Requirements for submitting variants and assertion criteria: - https://www.ncbi.nlm.nih.gov/clinvar/docs/submit/ - https://www.ncbi.nlm.nih.gov/clinvar/docs/assertion_criteria/

Main content	Discussion points
Nordic variant database	Added value compared to use of ClinVar - A Nordic database would be of high value if the labs submitted all their classified variants. - Differentiator: High quality Harmonization - Necessary anyway if submitting to ClinVar - Harmonization through common tools (ELLA, Scout) possible, but cannot be a requirement as this will exclude other contributors. - ELLA could be included as optional freeware - Tools must provide automatic submitting to avoid additional work operations Database content - Classification should be included as open information - Suggestion: need to be part of the Nordic network to exchange further information and prompt active curation (secure trust) - Entry requirements to secure high quality of database Access and access levels - Access should be via contribution (not to be counted) - Basic level: An open population frequency database - Second level: documentation on classification, by whom, when Annotation - annotation to be included in population frequency database? Adds complexity Technical and legal barriers to be explored and addressed

5.4 Actions

Table 11 Agreed actions

Theme	Action	Responsible	Deadline
Comparison of population-specific variants	Eidi to check if 38 non-ExAC OUS variants are present in SweGene	OUS/ Eidi	01.03.2017
Comparison of population-specific variants	Query SweFreq with own variants	OUS/ Svein Tore SciLifeLab/ Måns Rigshospitalet/ Wiktor, Ane	01.03.2017
ClinVar	Submit variants to ClinVar	OUS/ Svein Tore SciLifeLab/ Måns Rigshospitalet/ Wiktor, Ane	01.03.2017
ClinVar	Next WS: Comparison on experiences	All labs	April 2017

Sharing of variants - CASE - Population variant frequencies	Nordic variant frequency database CASE-development; 1-page document defining scope and approach for Nordic variant frequency database, to be used also for legal clarifications	DNV GL/ Sharm	21.02.2017
Sharing of variants – CASE - Curated single variants + phenotype	CASE-development; 1-pager on sharing of curated (single) variants, also to be used as basis for legal discussions	To be discussed further	
Legal	Approach Norwegian Health Directorate for clarifications on Nordic variant frequency database	DNV GL/ Guro	12.12.2016
Legal	Identify regulating authorities in NO, SE, DK from which to seek approval for sharing of variants	SciLifeLab/ Valtteri OUS/ Dag Rigshospitalet/ Morten	01.04.2017
Nordic variant frequency database	SweFreq database and tools sharing	SciLifeLab/ Henrik	01.04.2017
Nordic variant frequency database	Identify and access test dataset for aggregation	All labs OUS: SciLifeLab: NA Rigshospitalet:	01.04.2017
Nordic variant frequency database	Test SweFreq aggregation tool, feedback to developers	All labs OUS: SciLifeLab: Rigshospitalet:	01.04.2017
Benchmarking of variant interpretation	Benchmarking of variant interpretation per ACMG guidelines between OUS and Rigshospitalet	OUS/ Morten	01.04.2017

6 ENHANCE QUALITY OF DATA AND PROCESSES

During workshop 1 the clinical pipelines for exome / genome sequencing in the three clinical entities present were mapped to the below agreed process steps to identify similarities and differences in design and operations. Mapping included what is done, how is it done (software / hardware) and who does it (competence, institution) as summarized in the previous workshop summary report. As a first step in focusing on quality assurance, quality control steps and reference standards / guidelines used were identified as part of the mapping.

6.1 Benchmarking and variant interpretation

Agreed follow-ups from the previous workshop included initiation of a **technical benchmarking** exercise of sequence alignment and variant calling and variant annotation (green). There was also an expressed interest in sharing experiences on **variant interpretation** (red), including classification according to the ACMG[6] procedures / ranking of variants. Both these items were therefore on the agenda for this workshop. In addition, a framework for assessing data quality and data quality framework maturity was introduced.

[6] Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Richards, S. et al. Genetics in Medicine (2015) 17, 405-423.

Table 12 Enhance quality of data and processes

Agenda item	Main content	Discussion points
Benchmarking	Scope Methodology Status Proposed analysis Next steps	Benchmarking sample - NA12878 (part of GIAB[7] consortium) Benchmarking methodology - All labs to start with the same FASTQ files generated by SciLifeLab - Run through alignment, variant calling, annotation, filtering/ranking and reporting using standard production pipelines - Focus on standard bioinformatical annotation pipeline, not clinical annotation Reference genome - Analysis should be performed on the same reference genome (HG19), trying to perform on another and lifting over would introduce a lot of noise. Tools for variant calling - SciLifeLab use 3 callers; increase sensitivity to reduce false negatives and increase false positives. - OUS is evaluating which callers to use Variant annotation - SciLifeLab annotate on gene panel level. - In the benchmarking, it is more interesting to review annotations done in other labs than whole genome vcf files. Comparison basis - The files from the separate labs should be run against each other rather than a truth to compare variant calling outcomes. Benchmarking analysis - hap.py or bcbio.variation proposed as analysis tools Outcome - DNV GL will run analysis and report outcome Suggestions for future benchmarking exercises: - Classification to be included - Ranking of variants

[7] Genome In A Bottle Consortium; http://jimb.stanford.edu/giab

Agenda item		Main content	Discussion points
Harmonization of clinical interpretation of variants *What is needed to trust variant interpretations from other labs?*	OUS	Introduction of process - Interpretation process of samples and variants - Ad hoc strategies (research) - Weighting rules from ACMG guidelines[6] - OUS implementation of ACMG guidelines - Strength of ACMG codes - Documentation ELLA - OUS developed tool for variant interpretation; structured evaluation of annotation and references - Based on GenAP - Suggest relevant ACMG codes and classification - Replaces current documents with proper database - Currently in beta testing; production version expected Q1 2017	- Ethnicity not taken into consideration. - Using a gene panel of 500 genes may result in a list of 30 variants, where all are scored. Some may be easy & quick, other more laborious.
	Rigshospitalet - Department of Clinical Genetics	Introduction of process - Excel is used for variant score sheet based on ACMG guidelines. - Variant interpretation depends on how the variant has been filtered out. - Online version available from University of Maryland[8]	Discussion on classification and phenotype information - ACMG do not include phenotypes; disagreement in the larger genetic society on whether they should. - Lab reports should be objective and supporting evidence to treating physician, not a diagnosis. Phenotype should be considered in diagnosis, not in classification. - Need to differentiate between pathogenic and disease causing
	Rigshospitalet - Center for Genomic Medicine	Introduction of process - Germline variants from small Gene panels - Variant classification based on a combination of IARC and ACMG guidelines - Criteria for classification class 1-5 - Germline Variants from WES - Somatic variants from WES	

[8] http://www.medschool.umaryland.edu/Genetic_Variant_Interpretation_Tool1.html/

Agenda item	Main content	Discussion points
SciLifeLab	Introduction of process & SCOUT - Gene panels are defined; selected per case (e.g. inborn errors of metabolism) - Ranking developed together with CMMS; other clinics agree and keep the same system. - Ranking includes some 20 parameters - Ranking can be modified. - Rank_modelv1.18 - Next release will include SweFreq[3]	Ranking - Interest in learning more about the ranking process; to be put on the agenda for the next workshop Inheritance models - Rank model includes inheritance models. - Autosomal inherited recessive disorders can be filtered. - Gene of reduced penetrance: manually modified; affects the inheritance model. Learning system - Comments from interpreters are fed into a learning system (→ quality registry) - In the next update opportunity to add information about "the patient had this disease" will be included. Hosting - Hosted at University site as an intermediate solution while arranging for permanent solution with the hospital IT department Tools - Would like to link ELLA and SCOUT - Automation of classification using ELLA a long-term goal; ELLA now provides a suggestion and the user must confirm.

Agenda item		Main content	Discussion points
	Karolinska	Introduction of process - Utilization of ranked variants from SciLifeLab exemplified through case patient. - The clinician does not know what is wrong with the patients, and will go through the list according to strict guidelines. Will not do any filtering, going through the list of ranked variants takes 5 min. Each variant is easily accessed to check potential impact. - Two independent assessments are done and discussed at weekly meetings. Comments from assessors are available in Scout.	Authentication - Restricted IP address access - Accounts are applied for and handled - Two-factor authentication – google - Account closed when not used
	Information needed when exchanging curated variants	Information needed to trust interpretations from other labs	Whiteboard / post-its exercise, see Table 13 and Figure 2[9] below.
	Finding common ground		Phenotype indications - Introducing structured phenotype indications (e.g. based on Human Phenotype Ontology, HPO) from physicians a challenge - Too detailed phenotype indications could post a legal challenge with regards to privacy - GA4GH[10] Matchmaker Exchange[11] has interpreted which level of HPO terms are not sensitive. ACMG guidelines No global agreement, no framework in place

[9] For original mapping outcome, see appendix 3.

[10] http://genomicsandhealth.org/

[11] http://www.matchmakerexchange.org/

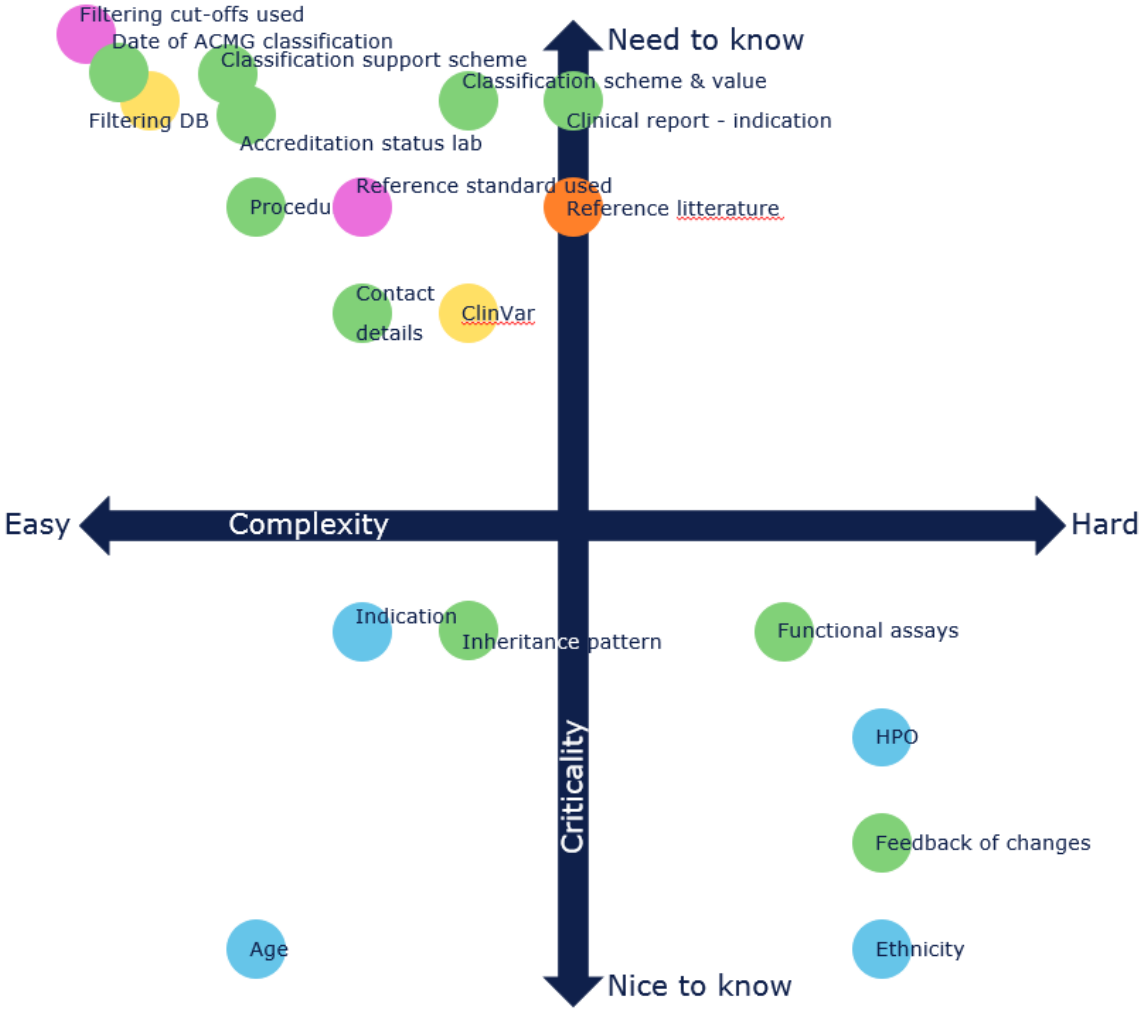
Agenda item	Main content	Discussion points
Data governance	DNV GL Data Quality framework – DNV GL Recommended Practice 0497 - Framework references - Data quality measurements - Organisational maturity assessment - Data maturity radar; levels and perspectives	Invitation to workshop participants to test framework - Pilot will be a semi assessment approach to test framework for genomic laboratories needs and value

Table 13 Categories of information needed to trust variant interpretation from other labs.

Category	Information item identified	Complexity Easy to provide (-4) Hard to provide (+4)	Criticality Nice to know (-4) Need to know (+4)
Other	Date of ACMG classification	-4	4
Other	Accreditation status of lab	-3	4
Other	Classification support scheme	-3	4
Other	Classification scheme and value	-1	4
Other	Clinical report relevant for indication	?	4
Filtering / prioritization	Filtering cut offs used	-4	4
External database	Filtering DB	-4	4
Other	Procedures	-3	3
Litterature	Reference; Classification or article, reference / literature + brief summary, comments	?	3
Filtering / prioritization	Reference standard used (ACMG/other)	-2	3
Other	Contact details of submitter	-2	2
External database	ClinVar – how interpreted	-1	2
Phenotype information	Indication / broad categories of referral reason	-2	-1
Other	Inheritance pattern	-1	-1
Other	Functional assays if done	2	-1
Phenotype information	HPO	3	-2
Other	Feedback of changes	3	-3
Phenotype information	Age	-3	-4
Phenotype information	Ethnicity	3	-4

? = disagreement between labs on complexity in providing information

Figure 2 Information needed to trust variant interpretation from other labs - criticality vs complexity in providing information[9]



6.2 Actions

Table 14 Agreed actions

Theme	Action	Responsible	Deadline
Benchmarking	Run analysis and share results	Rigshospitalet/ Wiktor	01.03.2017
Benchmarking	Share annotated outputs	OUS/ Svein Tore SciLifeLab/ Emma Rigshospitalet/ Wiktor, Ane	01.03.2017
Benchmarking	Compare variant calling, annotation results	DNV GL/ Sharm	01.03.2017
Benchmarking	Run gene panel-specific calling	DNV GL/ Sharm coordinate	01.03.2017
Benchmarking	Run and compare ranking results: true case	DNV GL/ Sharm coordinate	Backburner 01.03.2017
Harmonization of clinical interpretation of variants	Summary of variant interpretation pipeline for sharing	Karolinska/ Nicole	01.03.2017
Harmonization of clinical interpretation of variants	Sharing of SciLifeLab ranking process	SciLifeLab/ Måns/Henrik	Next workshop (agenda item) 01.03.2017
Data governance	Invitation to pilot DNV GL Data quality assessment framework;	DNV GL / Stephen	01.03.2017

7 NORDIC NETWORK FOR CLINICAL GENOMICS

7.1 Defining the Nordic added value

Table 15 Summary of discussion

Main content	Discussion points
General comments	- Clinical genomics is developing, as are several initiatives for efficient sharing of genomic data. - Interactions between academia and clinical setting part of the discussion. - It is premature to discuss the end solution - Need to understand frame conditions, but the long-term goals should be visionary and go beyond current restrictions - The added value of a Nordic cooperation must be identified, and cannot impose restrictions
Nordic common setting / strengths	- High standard public healthcare services - Unique ID numbers - eHealth status - Well-developed biobanks and registries - Regional healthcare organizations? - The Scandinavian model?
Nordic common challenges	- Nordic genetic variations - Small population sizes create a necessity to share - Development / implementation of national strategies for precision medicine - Tech transfer from research to clinic?
Nordic common interests – the added value of an extended network	- A common professional voice in influencing the frame conditions for implementation of precision medicine - We can share to learn and to lift our performances. Forum to discuss where to go and where not to go. - Solving problems together; sharing of methods, tools and experiences as important as sharing of data. - Opportunity to establish a learning system combining and looping back to contributors when differences of opinions - Build on Nordic common strengths - Nordic genetic variations
Think globally - act locally	- Global solutions tend to be built bottom up - Nordic laboratory for building solutions (MME, beacons ++)

7.2 Organization of further collaboration – defining the group

Table 16 Discussion

Main content	Discussion points
Group members	Cooperation will attract others, and although we should not recruit actively in this early phase, the collaboration should be open and eventually expanded Agreement that the collaboration is a network of organizations / departments
Formalisation of collaboration	Consensus that the collaboration should be formalized Benefits of formalization include opportunities to identify interesting projects and approach funding agencies, and for a consensus based common voice in the public debate. Rules for collaboration to be established with policy statements along the lines of the agreed goals.
Platform	- Sharepoint or Github sufficient? - Nordic cooperation forum on slack.com

7.3 Actions

Table 17 Agreed actions

Theme	Action	Responsible	Deadline	Comment
Collaboration	WS I summary report to be distributed to new members of network	DNV GL/ Guro	22.11.2016	Done
Collaboration	Distribute **Norwegian Strategy for Personalized Medicine** to members (SE, DK)	DNV GL/ Guro	31.12.2016	Link in this doc
Collaboration	Name contest, winner to be awarded next workshop	DNV GL/ Guro to coordinate	18.04.2017	
Collaboration	Propose principles for cooperation - Governance statues - Rules of collaboration incl. policy statements reflecting agreed goals - Confidentialities - Openness	DNV GL/ Vibeke	18.04.2017	
Collaboration	Value statement Nordic cooperation	DNV GL/ Guro	18.04.2017	
Platform	Nordic cooperation forum on slack.com	Rigshospitalet/ Wiktor	31.01.2016	Done; nordic-clinical-ngs.slack.com

8 NEXT WORKSHOP

Table 18 Agreed actions

Theme	Action	Responsible	Deadline
Next workshop	Identify venue in Copenhagen	DNV GL/ Guro	31.12.2016
Next workshop	Gather input and set optimal dates after Easter 2017	DNV GL/ Guro	12.12.2016
Next workshop	Gather input & develop draft agenda	DNV GL/ Guro	04.04.2017
Next workshop	Each group to define participants		

Table 19 Agenda items suggested for next WS - collected from Nov 2016 WS

Topic	Details	Responsible
Benchmarking	Report on outcome of annotation benchmarking	Sharm
Benchmarking	Variant interpretation per ACMG guidelines – OUS / Rigshospitalet. Harmonisation	OUS: Morten
Benchmarking	OSL to generate own FASTQ files from NA 12878?	OUS/Svein Tore
ClinVar	Experiences in submitting variants to ClinVar - Lab with most submitted variants to be rewarded!	Responsible to be nominated per lab OUS: SciLifeLab: Rigshospitalet:
Tools	ELLA – testing at SciLifeLab	Henrik / Måns / Robin
Clinical interpretation of variants	Introduction to SciLifeLab ranking process	Henrik / Måns / Robin
Nordic variant + phenotype database	Keep in view - scope and interest for developing this	Sharm
?	How to deal with structural variants	?
Harmonization of clinical interpretation of variants	Sharing of SciLifeLab ranking process	Måns/ Henrik

APPENDIX 1: WORKSHOP AGENDA

Monday 21 November 2016

Time	Topic	Responsible
11:00	Welcome & setting the stage	Dag & Vibeke
	Goals of collaboration	Vibeke
12:00	Lunch	
Ensuring legal compliance and engage with key stakeholders		
13:00	Intro	Guro
	Nordic national strategies - status and goals relevant for the Nordic cooperation.	Norway: Stephen Sweden: Valtteri Denmark: Morten
	Regulatory framework	Guro
	Relevant projects: BigMed. Others?	Vibeke
	Summary & next steps	Vibeke
Sharing of data		
14:30	Intro	Vibeke
	Sharing of variants	Sharm
	Sharing of tools and other resources: Gene panels	Måns
	Sharing of tools and other resources: Tools	Måns
	Sharing of tools and other resources: Procedures	Morten
	Summary & next steps	Vibeke
Ensuring data quality		
16:30	Intro	Vibeke
	Benchmarking	Sharm, Valtteri
17:40	End	

Tuesday 22 November 2016

Time	Topic	Responsible
8.30	Summary of day 1	
Ensuring data quality (2 of 2)		
9:00	Harmonisation of clinical interpretation of variants (+ procedures)	Morten, Eidi
Ca 10	- Short break	
12:00	Lunch	
13.00	Data governance	Stephen
13:05	Platform for sharing of tools and other resources (+ Gene panels & tools)	Måns
13.15	Breakout sessions: - Plan for sharing of Ella and Scout - Plan for developing case for sharing of curated variants (single)	
Vehicles for sharing		
Ca 14.00	- Short break, then cont. Freq. DB	
14.15	Breakout session: A - Variant frequency database B – Nordic Collaboration – what do we want to do and how do we continue?	Sharm (Svein Tore, Valtteri)
Wrapping up		
15:15	Summary of workshop – agree on next steps	Vibeke / Dag
16:00	End & departure	All

APPENDIX 2: LIST OF PARTICIPANTS

Table 20 List of participants

Organization	First name	Last name
SciLifeLab	Robin	Andeer
SciLifeLab	Emma	Sernstad
SciLifeLab	Måns	Magnusson
SciLifeLab	Henrik	Stranneheim
SciLifeLab	Valtteri	Wirta
Rigshospitalet	Wiktor	Mazin
Rigshospitalet	Karin	Wadt
Rigshospitalet	Morten	Dunø
Rigshospitalet	Ane	Yde Schmidt
OUS	Eidi	Nafstad
OUS	Knut Erik	Berge
OUS	Morten	Eike
OUS	Svein Tore	Seljebotn
OUS	Tony	Håndstad
OUS	Dag	Undlien
Karolinska	Nicole	Lesko
Karolinska	Anna	Wedell
DNV GL	Vibeke Binz	Vallevik
DNV GL	Brede	Børhaug
DNV GL	Guro Meldre	Pedersen
DNV GL	Sharmini	Alagaratnam
DNV GL	Stephen	McAdam

APPENDIX 3: HARMONIZATION OF CLINICAL INTERPRETATION OF VARIANTS

Figure 3 Mapping of information needed to trust interpretations from other labs.

APPENDIX 4: SLIDES USED DURING THE WORKSHOP

Files included

Nordic WS execution master

National strategies

- Norwegian Strategy for Personalised Medicine in Healthcare
- Sweden National status
- Denmark National status

OUS Existing sharing solutions

Harmonisation of clinical interpretation of variants

- OUS Clinical variant interpretation and ELLA
- Rigshospitalet Dep of clinical genetics – Interpretation
- Rigshospitalet Center for genomic medicine - interpretation

DNV GL Data Quality Assessment

DNV·GL

Nordic Clinical Genomics Data Sharing

Workshop Stockholm 21.-22. November 2016

Ungraded

Welcome

Ungraded

PARTICIPANT INSTITUTIONS

- Department of Medical Genetics, Oslo University Hospital, Oslo, Norway.
- SciLifeLab – Clinical Genomics Unit SLL, Sweden
- Karolinska Universitetssjukhuset – Centre for Inherited Metabolic Diseases (CMMS), Sweden
- Department of Clinical Genetics at The Juliane Marie Centre, Copenhagen University Hospital and the University of Copenhagen, Denmark.
- Center for Genomic Medicine at Diagnostic Centre, Copenhagen University Hospital and the University of Copenhagen, Denmark
- DNV GL

Oslo University Hospital

SciLifeLab

KAROLINSKA Universitetssjukhuset

REGION Rigshospitalet

DNV·GL

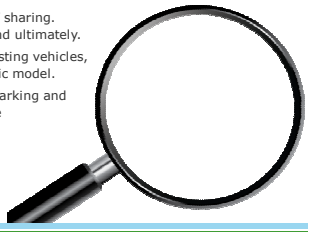
Ungraded

Practicalities

Ungraded

Goals for this workshop

- Agree on overall goal for the Nordic cooperation and how to get there
  - Ensure legal compliance and engage with key stakeholders: understand national strategies and current relevant projects.
  - Data to share: Review status of sharing. Define data to share, initially and ultimately.
  - Vehicles for sharing: review existing vehicles, identify gaps and propose Nordic model.
  - Ensure quality of data: Benchmarking and standardization of data to share

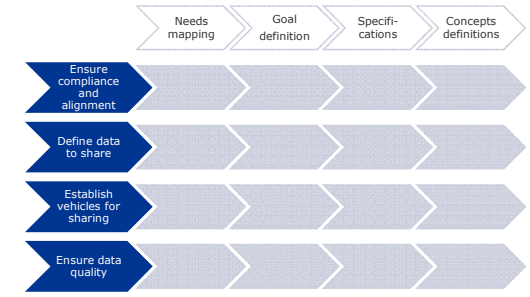


Ungraded

What are your expectations to this workshop?

Ungraded

Organization of workshop



Ungraded

7 DNV GL © 2017 21 November 2016 DNV·GL

Agenda Monday 21. Nov - plan

Time	Topic	Responsible
11:00	Welcome & setting the stage	Dag & Vibeke
	Goals of collaboration	Vibeke
12:00	Lunch	
Ensuring legal compliance and engage with key stakeholders		
13:00	Intro	Guro
	Nordic national strategies - status and goals relevant for the Nordic cooperation.	Norway: Stephen Sweden: Valtteri Denmark: Morten
	Regulatory framework	Guro
	Relevant projects: BigMed. Others?	Vibeke
	Summary & next steps	Vibeke
Sharing of data		
14:30	Intro	Vibeke
	Sharing of variants	Sharm
	Sharing of tools and other resources: Gene panels	Måns
	Sharing of tools and other resources: Tools	Måns
	Sharing of tools and other resources: Procedures	Morten
	Summary & next steps	Vibeke
Ensuring data quality		
16:30	Intro	Vibeke
	Benchmarking	Sharm, Valtteri
17:40	End	

8 DNV GL © 2017 21 November 2016 DNV·GL

Agenda Tuesday 22. Nov - plan

Time	Topic	Responsible
Ensuring data quality		
8:30	Harmonisation of clinical interpretation of variants	Morten, Eidi
	Data governance	Stephen
	Summary & next steps	Vibeke
Vehicles for sharing		
11:00	Intro	Vibeke
	Platform for sharing of tools and other resources	Måns
	Existing sharing initiatives – do they meet the needs?	Tony
12:00	Lunch	
13:00	Variant frequency database - Organization - Documentation and standardization of input	Sharm (Svein Tore, Valtteri)
	Summary & next steps	Vibeke
Wrapping up		
15:15	Summary of workshop – agree on next steps	Vibeke / Dag
16:00	End & departure	All

Ungraded

9 DNV GL © 2017 21 November 2016 DNV·GL

Agenda Monday 21. Nov - actual

Time	Topic	Responsible
11:00	Welcome & setting the stage	Dag & Vibeke
	Goals of collaboration	Vibeke
Ensuring legal compliance and engage with key stakeholders		
13:00	Intro	Guro
	Nordic national strategies - status and goals relevant for the Nordic cooperation.	Norway: Stephen Sweden: Valtteri Denmark: Morten
	Regulatory framework	Guro
	Relevant projects: BigMed. Others?	Vibeke
	Summary & next steps	Vibeke
Sharing of data		
14:30	Intro	Vibeke
	Sharing of variants	Sharm
	Sharing of tools and other resources: Gene panels	Måns
	Sharing of tools and other resources: Tools	Måns
	Sharing of tools and other resources: Procedures	Morten
	Existing sharing initiatives – do they meet the needs?	Tony
	Summary & next steps	Vibeke
Ensuring data quality (1 of 2)		
16:30	Intro	Vibeke
	Benchmarking	Sharm, Valtteri
17:40	End	

10 DNV GL © 2017 21 November 2016 DNV·GL

Agenda Tuesday 22. Nov - actual

Time	Topic	Responsible
8.30	Summary of day 1	
Ensuring data quality (2 of 2)		
9:00	Harmonisation of clinical interpretation of variants (+ procedures)	Morten, Eidi
Ca 10	- Short break	
12:00	Lunch	
13.00	Data governance	Stephen
13:05	Platform for sharing of tools and other resources (+ Gene panels & tools)	Måns
13.15	Breakout sessions: - Plan for sharing of Ella and Scout - Plan for developing case for sharing of curated variants (single)	
Vehicles for sharing		
Ca 14.00	- Short break, then cont. Freq. DB	
14.15	Breakout session: A - Variant frequency database B – Nordic Collaboration – what do we want to do and how do we continue?	Sharm (Svein Tore, Valtteri)
Wrapping up		
15:15	Summary of workshop – agree on next steps	Vibeke / Dag
16:00	End & departure	All

11 DNV GL © 2017 21 November 2016 DNV·GL

Workshop objectives – May 2016

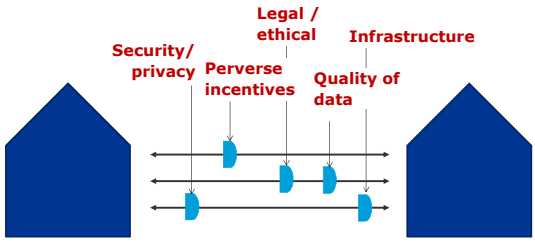
- Review **current clinical variant pipelines** in the three laboratories;
  - discuss common challenges
  - identify areas where standardisation/harmonisation could be beneficial.
- Identify what **specific data** would be valuable for laboratories to be able to share
  - short, medium and long term
  - current technical, legal and ethical barriers to sharing
- Discuss potential **models for future cooperation** and agree on next steps

Ungraded

12 DNV GL © 2017 21 November 2016 DNV·GL

Scope of workshop May 2016

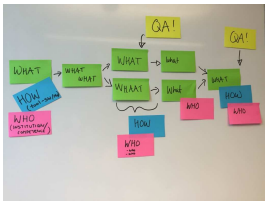
- How to facilitate sharing of data between labs – where to start?



Ungraded

13 DNV GL © 2017 21 November 2016 DNV·GL

MAPPING May 2016



- What – activities
- How – tools (hardware / software)
- Who – institution / competence

Ungraded

14 DNV GL © 2017 21 November 2016 DNV·GL

MAPPING OUTCOME

15 DNV GL © 2017 21 November 2016 DNV·GL

Topic: Organization of cooperation	Action responsible	Notes
Establish platform for sharing of information	SciLifeLab / Måns	- Github account? - Send information email to all participants
Funding	OUS / Dag	- Explore opportunities for funding
Stakeholder involvement	All	- Open approach; units at liberty to invite additional participants from own countries
Topic: Organization of follow-up workshop	**Action responsible**	**Notes**
Practical organization - Find date in November (doodle) - Practical planning of workshop II	DNV GL / Guro - connect with Anna	- November in Stockholm
Workshop agenda planning - coordinate input from all on agenda	DNV GL / Guro	- legal, data sharing, benchmarking - Separate work streams on biostatistical and legal issues
Topic: Legal	**Action responsible**	**Notes**
Understand legal basis and barriers for sharing data and establishing a common database	DNV GL / Guro (agreed with Dag)	- initiate and coordinate with OUS, Karolinska/SciLifeLab (Anna, Valtteri) and Rigshospitalet (Karin, Morten)
Topic: Data / information sharing **Goal:** - **First step: Establish basic level of sharing** - **Vision: Nordic sharing / database**	**Action responsible**	**Notes**
Sharing of gene panels	SciLifeLab / Måns	- Establish basis in repository - Check legal?
Share procedural docs	OUS / Morten	- Standard operating procedures - Sweden: Rank score definition files - Check legal?
Explore applicability of existing tools	OUS / Tony?	- Do existing tools serve the need for exchange of information in a curated database? o ClinVar, o Matchmaker / Beacon - Check legal?
Frequency database	OUS / Swpin Tore	- Variant frequencies / counting - Include research data? - Check legal?

16 21 November 2016

Topic: Benchmarking	Action responsible	Notes
Technical benchmarking exercise	SciLifeLab / Valtteri	- Easy first step to get started - Scope of benchmarking to be defined - Clinical genomics focus - Variant identification - Variant calling - Variant interpretation - Testing of 25-30 variants? - Benchmarking outcome to be presented at Workshop II.
Topic: Tools	**Action responsible**	**Notes**
Co-development of tools	SciLifeLab / Måns?	- To be followed up in Workshop II – separate stream for bioinformatics
Topic: Standardization	**Action responsible**	**Notes**
Concept document	DNV GL / Guro	- prepare draft, circulate
Process description	DNV GL / Guro	- Circulate for feedback - Process risk assessment → QC, QA, validation - Identify common standards / guideline references / etc. - What part of the process needs harmonizing to facilitate data sharing (based on data sharing exercise)

Ungraded

17 21 November 2016

Workshop rules

- Examples:
  - Phones and PCs?
  - Important phone calls → outside
  - SPA - Splng doesn'tt count
  - Parking lot
  - Confidentiality?

Ungraded

18 DNV GL © 2017 21 November 2016 DNV·GL

Goals of collaboration - draft

GOAL	Improved diagnosis and treatment of patients through responsible sharing of trustworthy data			
Sub-goal	Defining data to share	Establish vehicles for sharing	Ensure quality of data	Ensure legal compliance and engage with key stakeholders
Concepts/ solutions	Agree on data to share; initially and ultimately	Review current available solutions and agree on partner solution	Benchmarking, standardization of data to share	Comply with regulations and align with national/ regional/ international strategies

Ungraded

Revisited: Goals of collaboration V0.2

GOAL	We work together and learn from each other to lift our performance standards. We aim at responsible sharing of trustworthy data for improved diagnosis and treatment, and as a resource for research.			
Sub-goal	Defining data, tools and methods to share	Establish vehicles for sharing	Enhance quality of data and processes	Understand legal barriers and engage with key stakeholders

Ungraded

The Nordic added value – summary of discussions

- Nordic common setting / strengths
  - High standard public healthcare services
  - Unique ID numbers
  - eHealth status
  - Well developed biobanks and registries
  - Regional healthcare organizations?
  - The Scandinavian model?
- Nordic common challenges
  - Nordic genetic variations
  - Small population sizes creates a necessity to share
  - Development / implementation of national strategies for precision medicine
  - Tech transfer from research to clinic?
- Nordic common interests – the added value of an extended network
  - A common voice in influencing the frame conditions for implementation of precision medicine
  - We can share to learn and to lift our performances. Forum to discuss where to go and where not to go.
  - Build on Nordic common strengths
  - Nordic genetic variations
  - Solving problems together
- Think globally - act locally
  - Global solutions tend to be built bottom up
  - Nordic laboratory for building solutions (MME, beacons ++)

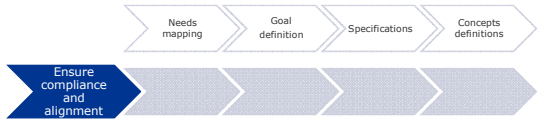
Ungraded

Ensuring legal compliance & engaging with main stakeholders

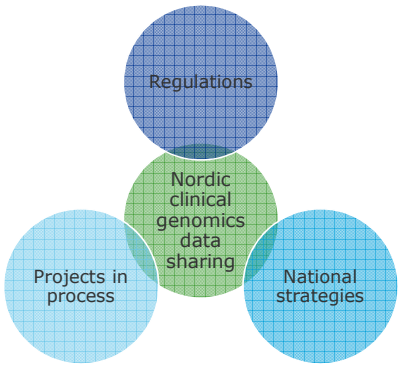
Ungraded

Ensuring legal compliance & engaging with main stakeholders

Ensuring legal compliance and engage with key stakeholders		
13:00	Intro	Guro
	Nordic national strategies - status and goals relevant for the Nordic cooperation.	Norway: Stephen Sweden: Valtteri Denmark: Morten
	Regulatory framework	Guro
	Relevant projects: BigMed. Others?	Vibeke
	Summary & next steps	Vibeke



Ungraded

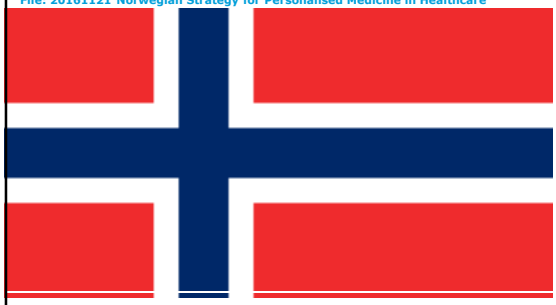


Ungraded

Ungraded

25 DNV GL © 2017 21 November 2016 DNV·GL

File: 20161121 Norwegian Strategy for Personalised Medicine in Healthcare



Ungraded

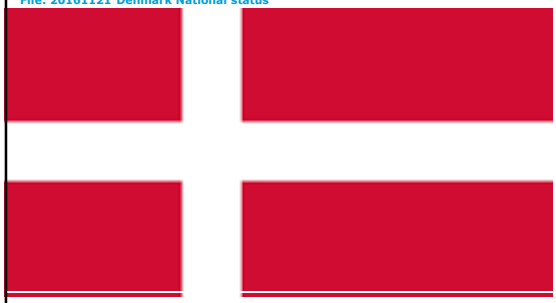
26 DNV GL © 2017 21 November 2016 DNV·GL

File: 20161121 Sweden National status

Ungraded

27 DNV GL © 2017 21 November 2016 DNV·GL

File: 20161121 Denmark National status



Ungraded

28 DNV GL © 2017 21 November 2016 DNV·GL

Regulatory framework

§

Topic: Legal	Action responsible	Notes
Understand legal basis and barriers for sharing data and establishing a common database	DNV GL / Guro (agreed with Dag)	initiate and coordinate with OUS, Karolinska/SciLifeLab (Anna, Valtteri) and Rigshospitalet (Karin, Morten)

- EU directive
- BigMed WP 5

Ungraded

29 DNV GL © 2017 21 November 2016 DNV·GL

EU General data protection directive / regulation

- Timeline
  - January 2012: The European Commission proposed a comprehensive **reform of data protection rules in the EU**.
  - 4 May 2016: the Regulation and the Directive published in the EU Official Journal in all the official languages.
  - While the **Regulation** will enter into force on 5 May 2016, it shall apply from **25 May 2018**.
  - The **Directive** enters into force on 5 May 2016 and EU Member States have to transpose it into their national law by **6 May 2018**.
- Objectives:
  - give citizens back control over of their personal data
  - simplify the regulatory environment for business
  - enable the Digital Single Market
- Provides legal basis for processing of personal information
  - With consent (explicit, well-defined, informed, voluntary)
  - Without consent (preconditions, purposes, safeguards)

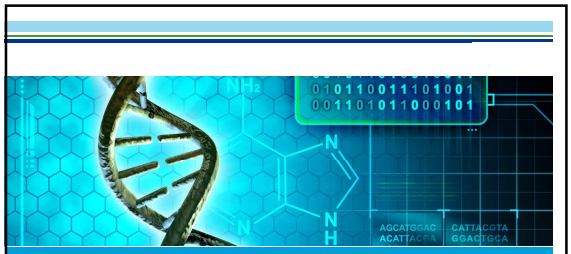
Ungraded

30 DNV GL © 2017 21 November 2016 DNV·GL

BigMed WP 5 - Legal and ethical issues and aspects

- **WP leader:** Anne-Kjersti Befring, UiO
- **Contributors:** E-helsedirektorat, Helsedirektoratet, OUS, USIT, All.
- **Goals**
  - Result goal: Demonstrate the **feasibility of sharing curated genomic data** from clinical labs in **two different countries**, identify potential legal **barriers** and suggest potential **solutions**.
  - Effect goal: **Necessary legal changes** to allow implementation of precision medicine are identified.
  - KPI: **BIGMED solution is legal** within expected near future legal framework in Norway.
- **Objectives:**
  - Ensure the project is aligned with the emerging **EU developments** addressing protection of personal data and provide combined technical and legal **input to national and international regulations under development**.
  - Engage interested parties to establish a legal and ethical team that will act as a "**centre of excellence**" to support the project to operate within acceptable legal and ethical boundaries.
  - Ensure external and internal visibility of the legal and ethical dimensions of the project.

Ungraded

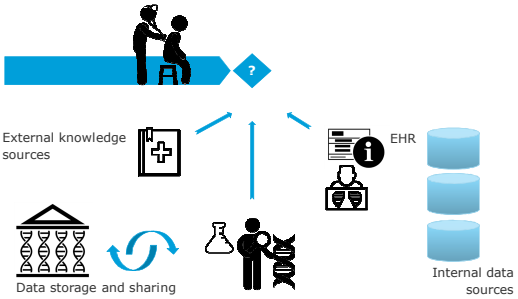


BIGMED

21 November 2016

Ungraded

Some bottlenecks for precision medicine?



Ungraded

Vision and goals for BIG data MEDical solution

VISION:
Lay the foundation for an ICT platform that addresses the analytic bottlenecks for the implementation of precision medicine and paves the way for novel big data analytics. The solutions will provide the patients with an optimized care which takes their unique individual characteristics into proper consideration.

E0 (1-4) - Enable use of Precision medicine information in diagnostics and targeted treatment

E1 Shared understanding of the needs in precision medicine among stakeholders

E2 Quality level of genomics information ensures practical clinical use

E3. Simplified access to and analysis of data For enabling precision medicine in clinical use

E4 Enable use of genomics information in diagnostics and treatment decisions through Enabling a basis for bigger data.

E5 - Necessary legal changes to allow implementation of precision medicine are identified

E6 - Increased public and political awareness of precision medicine

E7 - Innovation And new commercial activities around ICT in precision medicine in Norway

By precision medicine, we mean:

«...a medical model using characterisation of individuals' phenotypes and genotypes (e.g. molecular profiling, medical imaging, lifestyle data) for tailoring the right therapeutic strategy for the right person at the right time, and/or to determine the predisposition to disease and/or to deliver timely and targeted prevention."

Europarådet, 2015

Vision and goals for BIG data MEDical solution

VISION:
Lay the foundation for an ICT platform that addresses the analytic bottlenecks for the implementation of precision medicine and paves the way for novel big data analytics. The solutions will provide the patients with an optimized care which takes their unique individual characteristics into proper consideration.

E0 (1-4) - Enable use of Precision medicine information in diagnostics and targeted treatment

E1 Shared understanding of the needs in precision medicine among stakeholders

E2 Quality level of genomics information ensures practical clinical use

E3. Simplified access to and analysis of data For enabling precision medicine in clinical use

E4 Enable use of genomics information in diagnostics and treatment decisions through Enabling a basis for bigger data.

E5 - Necessary legal changes to allow implementation of precision medicine are identified

E6 - Increased public and political awareness of precision medicine

E7 - Innovation And new commercial activities around ICT in precision medicine in Norway

R1. Identify the requirements for precision medicine from the needs defined by the selected patient groups.

R2. Develop and demonstrate bioinformatics tools that will speed up the robust and scalable inclusion of genomic diagnostics into mainstream healthcare.

R3. Develop and demonstrate novel, beyond state of the art tools to extract actionable medical information from multiple unstructured and structured sources of

R4. Develop and demonstrate novel analytical ICT tools for integrative analysis of patient data from electronic health records, genomic datasets and phenotype data

R5. Demonstrate the feasibility of sharing curated genomic data from clinical labs in two different countries, identify potential legal barriers and suggest potential solutions.

R7. Conduct studies on the effect of the BIGMED solutions on patients' health and quality of life as well as for health care costs.

R8. Validate the genetic value of the tools developed on the four clinical use cases. They will also be tested on other diseases including the ability to identify soldiers at risk for specific injuries based on soldiers' medical cards and on a cohort of patients diagnosed with sarcoma.

R6. Establish the basis for commercialization of the project innovations and potential new commercial spin offs through our industrial partners.

R9. Reinforce a shared vision among the BIGMED partners and establish the consortium as an internationally recognized enabler of precision medicine.

Ungraded

Consortium partners

Juridisk rammeverk

Kravstillere

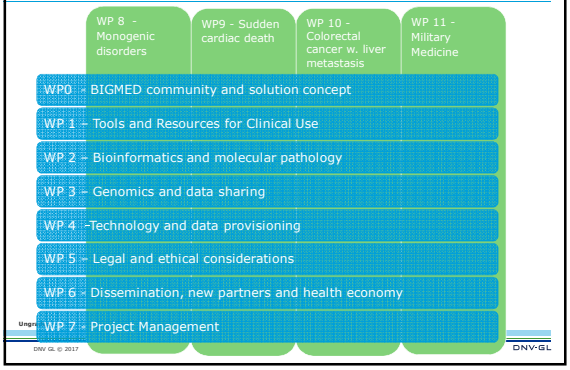
Pasienter

IKT-forskning

Industri og næringsliv

Ungraded

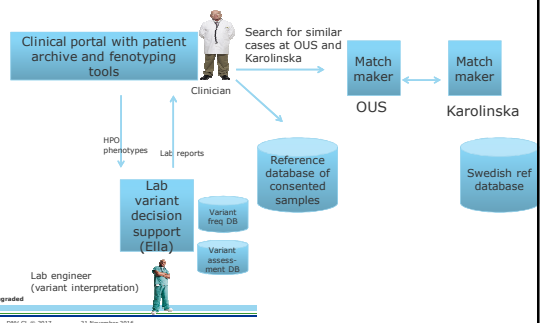
Workpackage structure



Workpackages – details

- WP0: BIGMED community and solution concept
  - Overview of precision medicine, vision and bottlenecks
  - Map needs and user stories, community for sharing
  - Overarching systems architecture
- WP2: Bioinformatics pipeline and molecular pathology
  - Develop clinical sequencing pipelines for clinical precision medicine.
  - Contribute to define, design and evaluate elements of a clinical decision support system for cancer to enable precision medicine.
- WP3: Genomics and data sharing
  - Make clinical genomics data available and actionable.
  - Investigate and build genomic databases and decision support tools that enable controlled use of structured data across different patient samples, facilitating identification of similar patient cases and thereby increasing the diagnostic sensitivity for rare diseases.
- WP5: Legal and ethical considerations
  - Ensure the project is aligned with the emerging EU developments addressing protection of personal data and provide combined technical and legal input to national and international regulations under development.

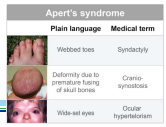
Conceptual overview of WP 3 - data sharing



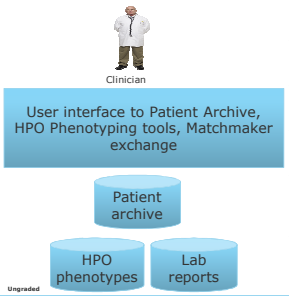
Doctors lack support for ordering genetic tests and describing cases

- For doctors, it can be difficult to know which test to order
  - Doctors of different competence levels order genetic tests (GPs, neurologists, to clinical geneticists)
  - Patients with rare syndromes in particular typically have a complex and somewhat heterogenous phenotype
- Important phenotype details are not captured and transferred to lab
  - Current paper-based requisition forms don't allow for detailed phenotype information

- **Structured phenotype data can guide test selection, steer the lab's variant search effort, and faciliate comparison of cases to identify similar patients**

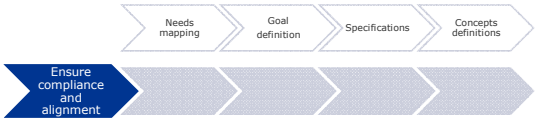


Clinical portal features



- Stores and retrieves genetics-specific data for patients and their families
- Generates HPO phenotype descriptions with support from available tools
- Offers decision support for selecting appropriate genetic test (e.g. gene lists)
- Displays variant interpretation reports from lab and ideally allows for communication "in context" with lab engineer/doctor
- Searches for similar patient cases in internal and external databases

Summary and next steps



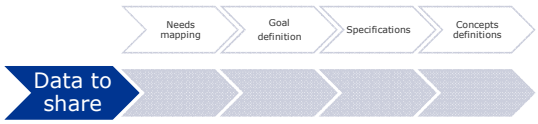
Defining data to share

Ungraded

43 DNV GL © 2017 21 November 2016 DNV·GL

Defining data to share

Data to share		
14:30	Intro	Vibeke
	Sharing of variants	Sharm
	Sharing of tools and other resources: Gene panels	Måns
	Sharing of tools and other resources: Tools	Måns
	Sharing of tools and other resources: Procedures	Morten
	Summary & next steps	Vibeke



Ungraded

44 DNV GL © 2017 21 November 2016 DNV·GL

Recap WS 1 – Data sharing

Priority #	Data	Adressed in ws 2
1	Population frequencies	Variants
2	Curated variant classification database	Variants
3	Genomes and patient phenotype	Variants
4	Genotype (full genomes) and patient phenotype database	Variants
5	Matchmaking through accurate and standardized phenotype descriptions	Vehicles
6	"Everything" – FastQ files with phenotypes	Variants
7	Benchmarking	Quality
8	Classification / ranking of variants , variant interpretation procedure (Application of ACMG)	Quality
9	Gene panels	Tools & other
10	QC procedure: coverage mapping, verification, etc.	Benchmarking?
11	Variants in hypernormal controls	?
12	Ability to query variant database by position	Variants

Ungraded

45 DNV GL © 2017 21 November 2016 DNV·GL

DNV·GL

Sharing of variants

Scope and ambition

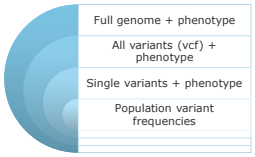
21 November 2016

Ungraded

46 DNV GL © 2017 SAFER, SMARTER, GREENER

Recap from Workshop I - and moving on

Priority #	Data	
1	Population frequencies	Variants
2	Curated variant classification database	Variants
3	Genomes and patient phenotype	Variants
4	Genotype (full genomes) and patient phenotype database	Variants
6	"Everything" – FastQ files with phenotypes	Variants



Brainstorming

- Why share?
- Existing solutions
- Nordic added value

Ungraded

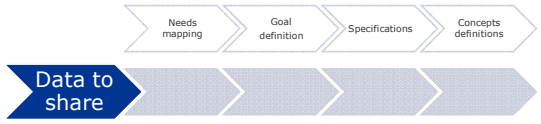
47 DNV GL © 2017 21 November 2016 DNV·GL

Sharing of
- gene panels
- tools
- procedures

Ungraded

48 DNV GL © 2017 21 November 2016 DNV·GL

Summary and next steps



Ungraded

49 DNV GL © 2017 21 November 2016 DNV·GL

Ensuring data quality to secure trust in shared data

Ungraded

50 DNV GL © 2017 21 November 2016 DNV·GL

MAPPING OUTCOME WS 1

51 DNV GL © 2017 21 November 2016 DNV·GL

WS 1 Outcome

- Agreed follow-up from WS 1
  - Technical benchmarking exercise
- Recap WS 1 – data sharing

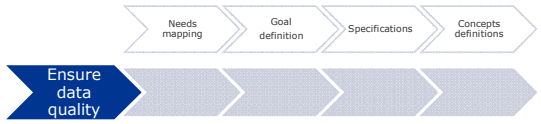
Priority #	Data	Adressed in ws 2
7	Benchmarking	Quality
8	Classification / ranking of variants , variant interpretation procedure (Application of ACMG)	Quality

Ungraded

52 DNV GL © 2017 21 November 2016 DNV·GL

Ensuring data quality

Ensuring data quality		
16:30	Intro	Vibeke
	Benchmarking	Sharm, Valtteri
Time	**Topic**	**Responsible**
Ensuring data quality		
8:30	Harmonisation of clinical interpretation of variants	Morten, Eidi
	Data governance	Stephen
	Summary & next steps	Vibeke



Ungraded

53 DNV GL © 2017 21 November 2016 DNV·GL

Benchmarking

Ungraded

54 DNV GL © 2017 21 November 2016 DNV·GL

DNV·GL

Benchmarking of NA12878

A first exercise

21 November 2016

Ungraded

Scope & methodology

- Scope
  - Sample: NA12878
  - Library prep: PCR-free, 75 ng input, Lucigen
  - Sequencing: HiSeq X, PE 2x150 bp, 1 sample per lane
  - Demultiplexing: 1 mismatch
  - Quantity of data: ca 450 M read pairs
- Methodology
  - All three labs to start with same FASTQ files generated by SciLife
  - Run through alignment, variant calling, annotation, filtering/ranking and reporting using standard production pipelines
  - Comparison of results by DNV GL, reporting

Ungraded

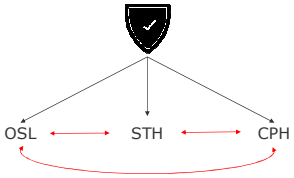
Status

	OSL	STO	CPH
Analysis status	Complete	Complete	Dataset downloaded, awaiting pipeline setup
Pipeline(s)	2 different: GATK v3 and Illumina	MIP v.3.0.7: Bwa, GATK v3, & 3 variant callers- GATKHaploypeCaller, Freebayes, Samtoolsmpileup Variant annotation VEP, snpeff	?
Output	VCF	VCF	?
	Genetic connective tissue disease gene panel, annotated	Metabolic disease gene panel VCF (ca 800 genes), annotated	?
	Whole genome annotated VCF	?	?

Ungraded

Proposed analysis

- Analysis tools: hap.py or bcbio.variation — Others?
- Truth dataset: GIAB vcf and bed — Others?
- Reference genome: hg19 — GChr38 for CPH? Liftover?



- Gene panel VCFs?
- Annotation? Whole genome?
- Ranking?

Ungraded

Next steps

- DK results: Wiktor
- Additional data from STH/OSL?
- Analysis: Sharm
- Reporting

Ungraded

Harmonisation of clinical interpretation of variants - what is needed to trust variant interpretations from other labs?

OUS intro + ELLA:
file: 20161121 OUS Clinical variant interpretation and ELLA
SciLifeLab intro + SCOUT:

Karolinska: summary to be provided

Rigshospitalet - Department of Clinical Genetics at The Juliane Marie Centre:
file: 20161121 Rigshospitalet Dep of clinical genetics - Interpretation
Rigshospitalet - Center for Genomic Medicine at Diagnostic Centre:
file: 20161121 Rigshospitalet Center for genomic medicine - interpretation

Ungraded

Data governance

see separate file

Ungraded

Establish vehicles for sharing

Ungraded

Recap WS 1 – data sharing

Priority #	Data	Adressed in ws 2
1	Population frequencies	Variants
2	Curated variant classification database	Variants
3	Genomes and patient phenotype	Variants
4	Genotype (full genomes) and patient phenotype database	Variants
5	**Matchmaking through accurate and standardized phenotype descriptions**	**Vehicles**
6	"Everything" – FastQ files with phenotypes	Variants
12	Ability to query variant database by position	Variants

Ungraded

Agenda Tuesday 22. Nov

Time	Topic	Responsible
8.30	Summary of day 1	
Ensuring data quality (2 of 2)		
9:00	Harmonisation of clinical interpretation of variants (+ procedures)	Morten, Eidi
Ca 10	- Short break	
11.00	Breakout sessions: A - Plan for sharing of Ella and Scout B - Plan for developing case for sharing of curated variants (single) + phenotype	
11.45	Data governance	Stephen
12:00	Lunch	
Vehicles for sharing		
13:00	Platform for sharing of tools and other resources (+ Gene panels & tools)	Måns
13:20	Breakout session: A - Variant frequency database B - Nordic Collaboration – what do we want to do and how do we continue?	Sharm (Svein Tore, Valtteri)
Ca 14	- Short break, then cont. Freq. DB	
Wrapping up		
15:00	Summary of workshop – agree on next steps	Vibeke / Dag
16:00	End & departure	All

Existing sharing initiatives - do they meet the needs?

File: 20161121 OUS Existing sharing solutions Stockholm Covered Monday.

Ungraded

Platform for sharing of tools and other resources

+ sharing of gene panels and tools

Ungraded

Breakout sessions:

A – Plan for sharing of Ella and Scout

B – Plan for developing case for sharing of curated variants (single)

Breakout sessions:

A - Variant frequency database
- organization
- documentation and standardization of input

B – Nordic Collaboration – what do we want to do and how do we continue?
Short term and long term goals
Format for collaboration
Opportunities for financing?

Variant frequency database
- organization
- documentation and standardization of input

DNV·GL

An anonymized Nordic frequency database?

Concept and potential application

21 November 2016

Why an anonymized Nordic frequency database

- Nordic-specific variants
  - 10% of coding variants in 400 normal genomes from Norwegian cancer patients absent in 1000G/ExAc
- Consolidating
  - Many isolated datasets scattered around Norway/Nordics
  - Norvariome, Norwegian Cancer Genome Consortium
  - SweFreq, Danish Reference Genome?
- Political
  - National Strategy for Personalized Medicine recommendation:
    A national anonymized frequency database to catalogue *normal* Norwegian variants for improving diagnostics
  - Proposed 2017 national budget allocates 5m NOK for this purpose

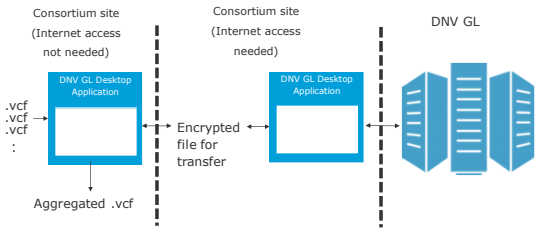
What do we want now – and later?

- Must-haves
  Allele frequency, observational count, variant annotation (specifically what?), quality indicators, technology, version, regional data, tracking of included samples (anonymized), site of origin
- Nice-to-haves
  Phenotype indicator (HPO? Hidden if under n=x?), genotype freq, gene panels membership, annotation conflict alerts, double individual alert
- Don't-wants
  Any identifiable data (person ID, DOB, etc)

System architecture

- File structure for sharing
  - Common tools for aggregating data: Beacon API?
  - Traceability of variants back to site of origin
  - N tested vs n seen: tech differences
  - Quality requirements
- Aggregated vs federated databases
- Update frequency: live vs. periodic
- Security
- Access and authorization

Possible solution



Query wishlist

Query	Return
Single variant	Frequency of this
Single variant	Frequencies of all variants in same gene
Full list of variants (3m?)	Frequencies for these, for ranking & filtering

National strategy Norway & implementation

Fra Høringsutkast nasjonal strategi for persontilpasset medisin i helsetjenesten, mars 2016

Anbefaling C1: Etablere nasjonalt anonymt frekvensregister for arvelige genvarianter

Bruk av resultater fra genetiske undersøkelser i pasientbehandling krever at man kan skille mellom normal genetisk variasjon og sykdomsgivende varianter. I arbeidet med å bestemme hvorvidt en gitt variant er sykdomsgivende eller ikke er det viktig å kunne sammenlikne funn hos en pasient med normalvariasjonen i befolkningen og eventuelt en tilhørende tidligere tolkning av tilsvarende funn. I dag benytter norske medisinskgenetiske avdelinger og laboratorier internasjonale databaser. Ulempen er at disse databasene ofte mangler normalvarianter som er spesielle for befolkningen i Norge. Det bør derfor etableres et nasjonalt, anonymt register over arvelige humane genvarianter. Registeret vil være et viktig klinisk verktøy, med kvalitetssikring av tjenesten som formål.

Fra Statsbudsjett 2017 for Helse og Omsorgsdepartementet

Regjeringen foreslår totalt 8 millioner til strategiarbeidet for persontilpasset medisin. Midlene skal gå til oppbygging og drift av en nasjonal, anonymisert database over genetiske varianter hos norske pasienter, samt etablering av et nasjonalt nettverk med oppbyggi
ng av regional, tverrfaglig kompetanse om persontilpasset medisin i alle helseregionene.

Summarize status and next steps

Expectations – which are met, which do we move to next steps?

- Fun
  - Ha det trävligt
  - There are no stupid questions – informal atmosphere
- Learn
  - How are other countries addressing legal challenges? (can we copy→paste?)
  - Learn how…., tools etc Scout
  - Learn how to use Scout
  - Learn about variant ranking in Scout
- Harmonising
  - Uniformity on which variants to report
  - Inspire to use ACMG criteria in var. int.
- First level sharing
  - Get our hands on Ella
- Fq. Database
  - Understand user needs
  - Develop consensus around technical req. + needs for anonymous freq. db
  - Sharing var. frequencies
  - Understand functional expectations
  - Identify misconceptions
  - MVP
- Cur. Variants:
  - One step further to sharing curated data
  - Get closer to nordic database of variants
- Cooperation
  - 1 step further – action oriented
  - Explore opportunities for nordic forum
  - Define Nordic added value
  - Nordic collaborateion forum slack (?)
  - Find concrete ways of sharing some of the easier things

Revisited: Goals of collaboration V0.2

GOAL

We work together and learn from each other to lift our performance standards.
We aim at responsible sharing of trustworthy data for improved diagnosis and treatment, and as a resource for research.

Sub-goal

- Defining data, tools and methods to share
- Establish vehicles for sharing
- Enhance quality of data and processes
- Understand legal barriers and engage with key stakeholders

Ungraded

Next steps for the Nordic network?

Formalize cooperation?
Continue with workshops?
…?

Ungraded

Thank you!

www.dnvgl.com

SAFER, SMARTER, GREENER

Ungraded

Strategy for Personalised Medicine in Healthcare

- Task given by Ministry of Health and Care Services
- Implementation in public healthcare systems
- 2017 - 2021
- No specific funding
- Under current legislation

Process

- Assignment from the Ministry of Health and Care Services
- Developed in collaboration with health professionals, researchers, NGOs, pharmaceutical industry
- Published June 2016
- 2017-2021 Perspective:
- The Directorate is responsible for implementation

More targeted and personalised healthcare

1. **Our healthcare service provide high-quality and relevant information and guidance on personalised medicine**
2. **Our healthcare service will implement personalised medicine as part of its services, and organisation of services and building of infrastructure will take place in nationally coordinated processes**
3. **Our healthcare service will contribute to research and innovation in the field of personalised medicine, both nationally and internationally**

Strategic Areas

A. **Expertise and information**
B. **Quality and academic and clinical development**
C. **Health registries**
D. **Information and communication technology (ICT)**
E. **Research and innovation**

A Expertise and information

A1: Include personalised medicine as a topic in relevant educations

A2: Establish a national network of regional resource centers for personalised medicine

A3: Develop national competence standards for genetic counselling

A4: Information for the public

B Quality and academic and clinical development

B1: Develop action plans

B2: Issue normative documents and standards for the clinical use of high-throughput technologies and genome-wide analyses

B3: Analyse the need for and eventually develop quality standards

C Health registries

C1: Establish a national and anonymous genetic variant database

C2: Further develop the Norwegian Cancer Registry to include more information on cancer genome variants

C3: Consider whether there is a need to include genome tests in the National registry on communicable diseases, and thus, the need for further developments on this registry

C4: Further investigations of the possibilities to establish a national system for storage and processing of raw data/medical information from clinical genome tests and analyses, both for healthcare purposes and for research.

D Information and communication technology (ICT)

D1: Further investigations of the possibilities to establish a national system for storage and processing of raw data/medical information from clinical genome tests and analyses, both for healthcare purposes and for research.

D2: To develop functionalities for handling of "personalised medicine"/large-scale data through the Electronic Patient Record (EPR).

E: Research and Innovation

- Action plan for PM across RCN programmes and the specialised health services:
  - Bedrehelse
  - Helsevel
  - Behandling
  - Helse-omsorg 21
  - BIGMED
- Develop patient pathways that integrate clinical treatment and research

The way forward

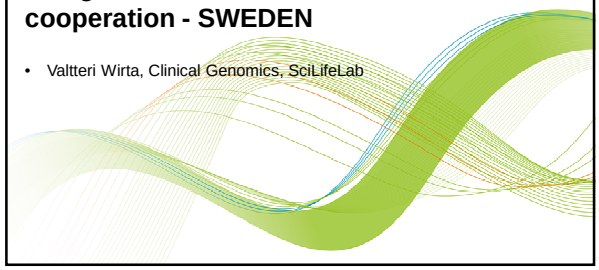
- The Health Directorates role:
  - National leadership and coordination
  - Harmonisation and guidance
- Follow-up in the context of relevant white papers and processes:
  - White paper on prioritisation
  - Review of the biotechnology law
  - National budget (St. prp. 1)
- National and International collaboration
  - Nordic cooperation?

SciLifeLab

2016-11-20

Nordic national strategies - status and goals relevant for the Nordic cooperation - SWEDEN

- Valtteri Wirta, Clinical Genomics, SciLifeLab



Clinical Genomics current strategy

SciLifeLab

- Focus on WGS for rare disease diagnostics
  - Improve current offering
  - Variant sharing
- Investigating the possibility to enter cancer genomics arena
- Formalising collaboration with the Karolinska University hospital ('joint unit')
- International (European) networking to identify potentials for collaborations
- SciLifeLab
  - SweFreq

2

National strategy for genomic medicine

SciLifeLab

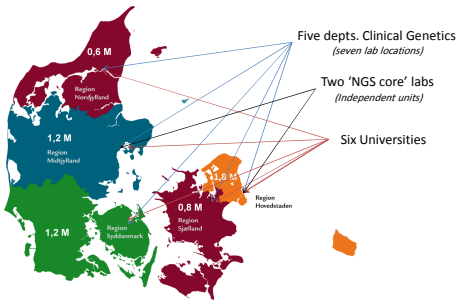
- No national strategy for genomic medicine in place
- Key players
  - Healthcare regions
  - Next-Generation Diagnostics platform at SciLifeLab
  - Certain research funders (eg Barncancerfonden)
- No national or regional political discussions ongoing (afaik)
- Within NGD and individuals associated with healthcare regions there is an early phase discussion
- Focus areas
  - Rare disease diagnostics – WGS
  - Solid tumours – mid size panels (300-500 genes)
  - Hematological malignancies – mid size panels (as above)

3

DK national strategy for NGS

"Precision medicine"

NGS landscape in Denmark



Copenhagen area

Health sector

Rigshospitalet
- Clinical Genetics
- Genomic Medicine
- Pathology
- Immunology
- Cardiology
- ao....

Herlev Hosp.
- Pathology

Hvidovre Hosp.
- Biochemistry

Statens Serum Institut , SSI

Research / education

- CHP Uni, Panum
- CHP Uni, Geogenetics
- DTU
- Psychiatric (St Hans)

Others

- Forensic institute
- **BGI**
- Exicon

Numerous stakeholders for NGS based analysis

GENOME denmark

The GenomeDenmark platform

A public initiative funded by a substantial grant (86 M. DKr) from the Innovation Foundation in 2011

The consortium behind consists of four Danish universities (KU, AU, DTU and AAU), two hospitals (Herlev and Vendsyssel) and two private firms (Bavarian Nordic and BGI-Europe).

Two Demonstration projects

Project one
- To develop new methods of pre-treatment for complex samples with traces of microorganisms being identified using sequencing
- To identify unknown microorganisms in cancer samples
- To identify new microorganisms associated with cancer development
- To select and describe microorganisms that can lead to the development of new diagnosing methods and anti-cancer vaccines

Project two
- To establish a high quality Danish reference genome (via 50 high quality genome trios).
- To generate knowledge that can support the development of personalized treatment, based on genomic information, in the health care system.
- To generate knowledge that can be applied to the Danish pharmaceutical and food industries

Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios.
Nat Commun. 2015 Jan 19;6:5969

Stakeholders from the 4 universities and representatives from Danske Regioner
(...but no clinical representatives)

-> Strategy paper outlining visions for a large scale DK sequencing project

Personlig Medicin og Individualiseret Behandling

Oplæg til en samlet dansk indsats
21 Juni 2015

DANISH e-INFRASTRUCTURE COOPERATION

Computerome (Nov 14)
16.048 CPU cores with 92 terabyte RAM, 4,5 PetaBytes storage

SDU Supercomputer (March 15, May16)
14.016 CPU cores, 766 teraflop/s

AAU (?? 2016/17)
??

...We just need decisions and funding....

DNV·GL

Existing sharing initiatives – do they meet the needs?

Ungraded

Agenda

- Frequency databases
- Variant classification and case-level databases
- Federated local databases / GA4GH
  - The Beacon Network
  - Matchmaker Exchange
- Plans for the BigMed project

Ungraded

Allele frequency databases

- Helps us to identify and filter out variants that occur commonly in a population
- Well known large public initiatives:
  - **ESP:** 6,503 exomes
  - **1000 Genomes:** Exome and WGS from 2,504 individuals
  - **ExAC**: 60,706 unrelated individuals exome-sequenced as part of various disease-specific and population genetic studies.
  - **gnomAD:** The next "ExAC" with 126,216 exome sequences and 15,136 whole-genome sequences

Ungraded

Example demonstrating usefullness of allele frequency databases

- Movement disorder panel (240 genes) results in **1130** variants
- After ExAC filtering (include <1% or NA): 458 variants
- 353 of these 458 are found in >5% of samples in our *internal* freq database
  - 317 of these 353 are in intronic regions (not covered by ExAC)
- **38** remain after filtering (mostly in UTRs not covered by ExAC)

Ungraded

No standard solution for local freqDB and sharing exist

- Most labs keep an internal database of allele counts
- But no standard exists, in-house developed solutions are not designed for sharing
- Contributing raw clinical (sensitive) data to e.g. ExAC pipeline is problematic
- A need for a Nordic frequency database (Nordic variants)?
- Valuable improvement?
  - Phenotype data
  - Data from healthy individuals
  - More ethnic diversity

Ungraded

Variant and case-level classification databases

- Capture knowledge from variant interpretation work and published litterature
- Examples of public and commercial alternatives:
  - HGMD (Univ. of Cardiff / Qiagen)
  - LOVD, BIC, IARC, UMD/BRCA Share, …
  - Decipher (Sanger Institute)
  - Clinvar (NCBI)

- The vast majority of variant classifications are still stored in local databases and not shared between labs

Ungraded

Clinvar – Our best alternative?

- **Quality**: Submission of variants requires submitting evidence for classification, as well as a publication or written statement of variant interpretation procedures. But no guarantee of quality of content. Classification discordances are transparent. A system for grading trustability/review status of a classification (★).
- **Ease of use and availability**: Free and public, but laborious to submit data
- **Completeness/offering**: Clinvar has gained momentum and will most likely be the biggest variant classification database (Nov 2016: 240,150 records submitted, 149,957 with assertion criteria).

Ungraded

7 DNV GL © 2017 21 November 2016 DNV·GL

The Global Alliance for Genomics and Health

Global Alliance for Genomics & Health

Beacon Project »
Tests the willingness of international sites to share genetic data in the simplest of all technical contexts.

BRCA Challenge »
Aims to advance understanding of the genetic basis of breast and other cancers using data from around the world.

Cancer Gene Trust »
The Cancer Gene Trust is an online network for sharing somatic cancer genomic and clinical data from around the world.

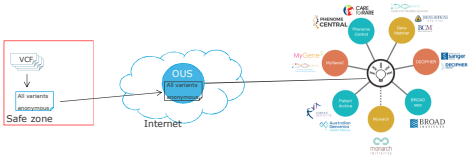
Matchmaker Exchange »
A federated network of rare disease data sets, which allows researchers to locate data on rare phenotypes or genotypes of interest.

Ungraded

8 DNV GL © 2017 21 November 2016 DNV·GL

Beacon – "Is this variant known to you"?

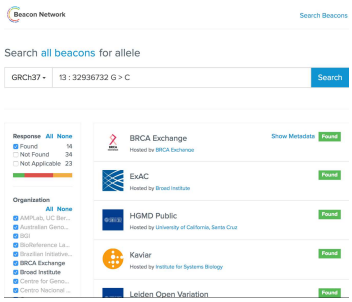
- A beacon is an *open* web service that any institution can implement and answers questions of the form "Do you have information about the following mutation?"
- The response is a simple "Yes" or "No", among potentially more information
- The Beacon *Network* is a search engine across the world's public beacons. It enables global discovery of genetic mutations, federated across a large and growing network of shared genetic datasets.



Ungraded

9 DNV GL © 2017 21 November 2016 DNV·GL

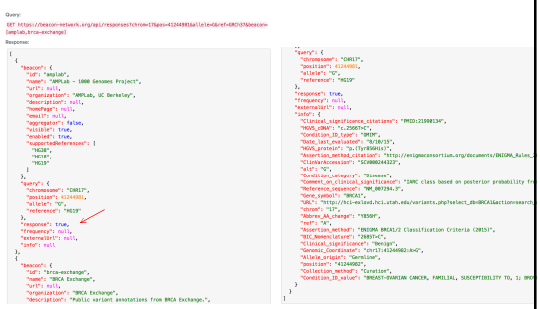
Querying the Beacon network manually



Ungraded

10 DNV GL © 2017 21 November 2016 DNV·GL

A computer can query the Beacon network via an API

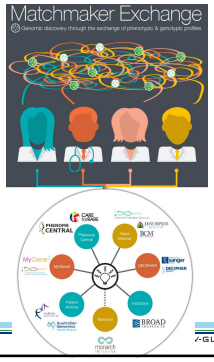


Ungraded

11 DNV GL © 2017 21 November 2016 DNV·GL

Matchmaker Exchange - Platform for Rare Disease Gene Discovery

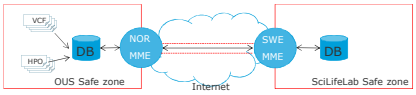
- A federated platform (exchange) to facilitate the identification of cases with similar phenotypic and genotypic profiles (matchmaking) through a standardized application programming interface (API) and procedural conventions.
- Supported by IRDRC, ClinGen, GA4GH
- MME enables queries containing phenotype (HPO) and gene/variant info to be sent to other MME services that then evaluate and return info about any matching similar cases



Ungraded

12 DNV GL © 2017 21 November 2016 DNV·GL

Query



Mandatory information in query:
- Case ID
- Submitter information
- Candidate gene/s and/or - phenotypes (HPO)

Additional information:
- Age of onset
- Mode of inheritance
- Condition name (OMIM, Orpha)
- Chromosome (and region)
- Variant type (frameshift, missense, etc.)

Ungraded

13 DNV GL © 2017 21 November 2016 DNV·GL

How to query and respond

```
{
  "patient" : {
    "id" : <identifier>,
    "label" : <identifier>,

    "contact" : {
      "name" : "Full Name",
      "institution" : "Contact Institution",
      "href" : <URL>
    },

    "species" : <NCBI taxon identifier>,
    "sex" : "FEMALE"|"MALE"|"OTHER"|"MIXED_SAMPLE"|"NOT_APPLICABLE",
    "ageOfOnset" : <HPO code>,
    "inheritanceMode" : <HPO code>,

    "disorders" : [
      {
        "id" : "MIM:######"|"Orphanet:#####"|…
      },
      …
    ],
    "features" : [
      {
        "id" : <HPO code>,
        "observed" : "yes"|"no",
        "ageOfOnset" : "…"
      },
      …
    ],
    "genomicFeatures" : [
      {
        "gene" : {
          "id" : <gene symbol>|<ensembl gene ID>|<entrez gene ID>
        },
        "variant" : {
          "assembly" : "NCBI36"|"GRCh37.p13"|"GRCh38.p1"|…,
          "referenceName" : "1"|"2"|…|"X"|"Y",
          "start" : <number>,
          "end" : <number>,
          "referenceBases" : "A"|"ACG"|…,
          "alternateBases" : "A"|"ACG"|…
        },
        "zygosity" : <number>,
        "type" : {
          "id" : <SO code>,
          "label" : "STOPGAIN"
        }
```

```
{
  "results" : [
    {
      "score" : {
        "patient" : <number>
      },
      "patient" : {…},
    },
    …
  ]
}
```

14 DNV GL © 2017 21 November 2016 DNV·GL

Plan for BigMed

- Make a report on functional and technical requirements of genomic databases
- Exchange frequency data between Nordic countries
- Build an OUS (Norwegian?) reference database of consented samples
- Implement a Beacon service
- Implement a PoC for Matchmaker Exchange

Ungraded

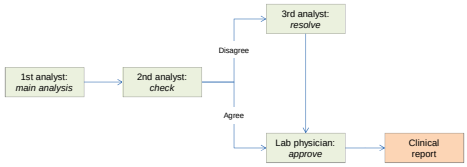
15 DNV GL © 2017 21 November 2016 DNV·GL

Clinical variant interpretation at Oslo University Hospital

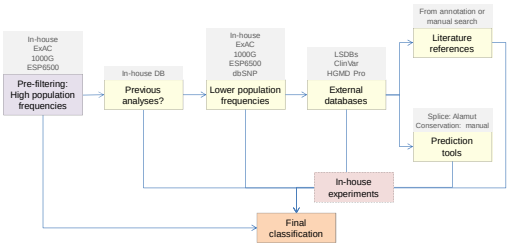
Data sharing WS 22.11.2016

Morten C. Eike | Post Doc

Interpretation of samples

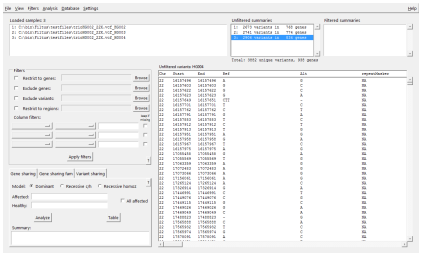


Interpretation of variants



Ad hoc strategies (research)

- Software: FILTUS
  github.com/magnusdv/filtus | folk.uio.no/magnusv/filtus.html
- Parts of source code also implemented in HTS trio pipeline: de novo and recessive



-rules from ACMG

Classification	Rule	Colour
Pathogenic	(i) 1 Very strong (PVS1) AND (a) ≥1 Strong (PS1–PS4) OR (b) ≥2 Moderate (PM1–PM6) OR (c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) OR (d) ≥2 Supporting (PP1–PP5) (ii) ≥2 Strong (PS1–PS4) OR (iii) 1 Strong (PS1–PS4) AND (a) ≥3 Moderate (PM1–PM6) OR (b) 2 Moderate (PM1–PM6) AND ≥2 Supporting (PP1–PP5) OR (c)1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5)	
Likely pathogenic	(i) 1 Very strong (PVS1) AND 1 moderate (PM1–PM6) OR (ii) 1 Strong (PS1–PS4) AND 1–2 moderate (PM1–PM6) OR (iii)1 Strong (PS1–PS4) AND ≥2 supporting (PP1–PP5) OR (iv) ≥3 Moderate (PM1–PM6) OR (v) 2 Moderate (PM1–PM6) AND ≥2 supporting (PP1–PP5) OR (vi)1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5)	
Benign	(i) 1 Stand-alone (BA1) OR (ii) ≥2 Strong (BS1–BS4)	
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) OR (ii) ≥2 Supporting (BP1–BP7)	
Uncertain significance	(i) Other criteria shown above are not met OR (ii) the criteria for benign and pathogenic are contradictory	

GENETICS in MEDICINE | Volume 17 | Number 5 | May 2015

ACMG at OUH

Goal:

Better reproducibility, standardisation across units

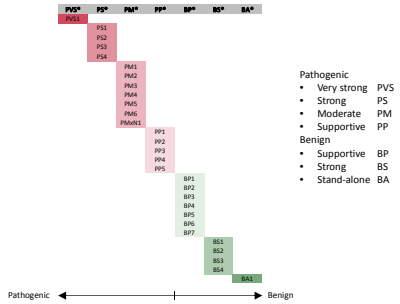
Implementation:

Working group

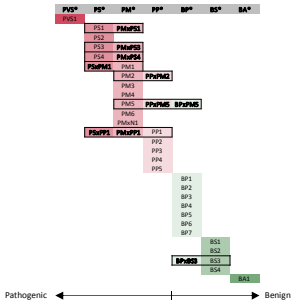
Status:

Guidelines adopted with some modification/ clarification

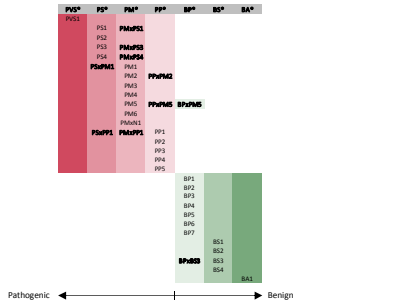
Strength of ACMG codes



Strength of ACMG codes



Strength of ACMG codes

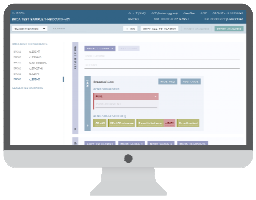


Documentation

- Variant evaluation document
  - Classification (1-5)
  - ACMG codes
  - Free text summary
  - Supporting observations (structured)
- Excel “database”
  - Variants with sample, classification and date of analysis

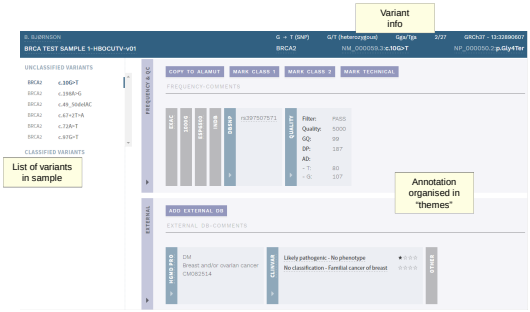
ELLA

Variant interpretation with *ella*

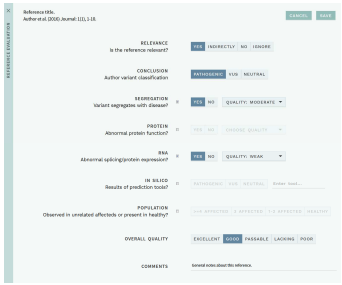


- Structured evaluation of annotation and references
- Suggest relevant ACMG codes and classification
- Replaces current documents with proper database

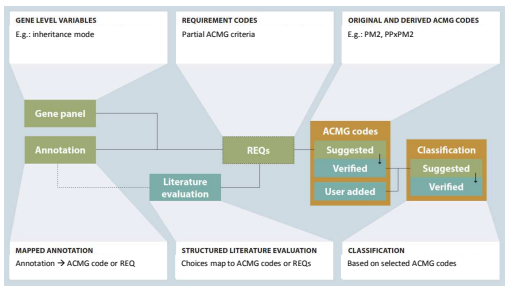
Annotation



Reference evaluation

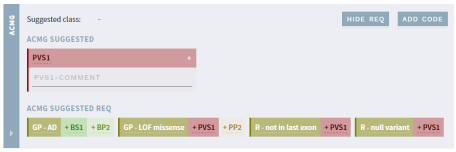


ACMG codes and classification



ACMG in user interface

- Suggested REQ + ACMG code it belongs to



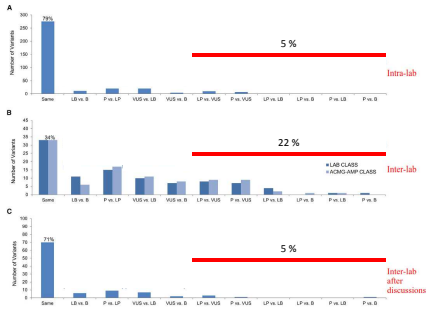
- ACMG code only suggested if all REQs are present

Status

- Currently in beta testing
- Production version early 2017

BACKUPS

Performance of ACMG guidelines



From Amendola et al. 2016, AJHG 98:1067-76

Variant classification

DK Rigshospitalet

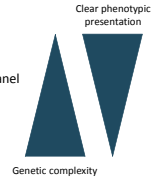
Adhere to the ACMG guidelines

Variant score sheet

http://medschool.umaryland.edu/Genetic_Variant_Interpretation_Tool1.html

Variant classification

- Single gene request
- Small focused gene panel
- Large gene panel
- WES/WGS



Is there a mutation in that gene?

Could a mutation in this gene explain the phenotype?

WES swaps burden of proof….

Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases

Nature Biotechnology **34**, 531–538 (2016) | doi:10.1038/nbt.3514

Received 29 July 2015 | Accepted 12 February 2016 | Published online 11 April 2016

Variant classification

Adhere to the ACMG guidelines

Panel

ad hoc discussions

WES

WES team meetings every 2. week
Medical Geneticist
Clinical Laboratory Geneticist

Always relate the genetic finding to the clinical presentation

Segregation analysis if possible

Flagged in in-house db as important
Recorded in Alamut
Criteria's stated in the clinical report

First national meeting in variant classification held in Århus September

Second scheduled for March 17 in Cph.

Variant interpretation –
Center for Genomic Medicine, Rigshospitalet

Nordic Clinical Genomics Data Sharing Workshop
Stockholm 21-22 november 2016

Germline variants from small Gene panels

- Variant (including CNV) calling in SeqPilot (JSI)
- Variant assesment always include a description of the variant.
  - Database (BIC, LOVD ect.)
  - Literature search
  - Population frequency (1000 Genomens, ESP, ExAC)
  - *In silico* data (Alamut)
  - Functional data (literature, in house, collaborations)
- Variant classification based on a combination of IARC and ACMG guidelines

Class 5

- Nonsense and frameshift mutations destrying expression of known functionalt domain
- Variants which in *in vitro* assays has shown to produce a transcript with a premature stop
- Destruction of the start site (no new in frame start site used)
- Copy number deletions creating a frameshift or deletion of a functionalt domain.
- Copy number duplications which creates frameshifts
- Variants with a probability of pathogenicity of >0.99 using multifactorial likelihood models

Class 4

- Missense variants resulting in an amino acid change which has be classified as class 5 by another variant.
- Small in-frame variants which result in deletion of amino acids classified as class 5.
- IVS -/+2 variants which have not been functionally investigated and do not result in a small in-frame ins/del or a natural in-frame transcript
- Variants with a probability of pathogenicity of 0.95-0.99 using multifactorial likelihood models

Class 3

- Variants lacking clinical and/or molecular evidence
- Variants with contradicting evidence
- Variants which *in vitro* have been found to result in natural transcripts not lacking any known functional domains.
- Variants with a probability of pathogenicity of 0.05-0.95 using multifactorial likelihood models

Class 2

- Missense variants resulting in an amino acid change which has be classified as class 1 by another variant and have *in vitro* been found to have normal splicing.
- Synonyms variants with low bioinformatic prediction of discupting normal splicing
- Missense variants, small in-frame variants intron variants with low bioinformatic probability of disrupting normal splicing and an allel frequency of 0.1-1% in a control group
- Variants with a probability of pathogenicity of 0.01-0.05 using multifactorial likelihood models

Class 1

- Variants found i ≥1% of control/reference group
- Variants with a probability of pathogenicity of <0.01 using multifactorial likelihood models

Germline Variants from WES

- CLC genomics workbench for variant calling
- VCF > Ingenuity – variant filtering
- Use germline guidelines for variant reporting

Somatic variants (WES)

- CLC genomics workbench for variant calling
- VCF > Ingenuity – variant filtering
- Main focus on class 4 and 5
- Treatment relevant mutations (hotspots)

DNV·GL

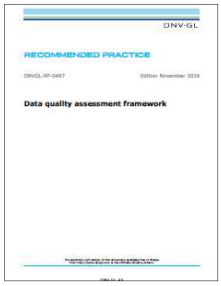
DNV GL Data Quality Framework

DNV GL Recommended Practice 0497

Ungraded

DNV GL Data Quality Assessment Framework

- Currently being developed together with customers that recognize this area is a business critical issue
- Designed to build trust between partners that depend on sharing data
- Two parts:
  - Generic tool for data quality assessment
  - Data quality framework maturity evaluation



Ungraded

Sources of inspiration for the Framework

DNV GL Data quality framework topics

1. Data Governance
2. Data quality processes
3. Data requirement definitions
4. Data quality metrics and dimensions
5. Data quality performance measurement
6. Architecture, tools and technologies
7. Data Standards

DNV GL Experience

ISO 8000 - 8

Data quality framework by David Loshin

DMM from CMMI

Ungraded

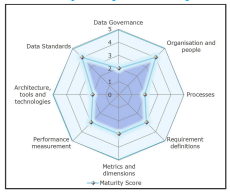
Data quality measurements – Organisational maturity assessment

Data

Data quality measurements

Organisation

Data quality maturity



Data quality results shown was collected as part of a monitoring and control system of System 27-Gas Export System on board of an offshore production platform.

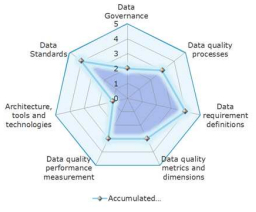
Ungraded

DATA MATURITY RADAR

Example of visualization of a data quality process maturity assessment is shown below.

Framework element	Score
Data Governance	2
Data quality processes	3
Data requirement definitions	4
Data quality metrics and dimensions	3
Data quality performance measurement	3
Architecture, tools and technologies	1
Data Standards	4

Accumulated Score with Customer 1



Ungraded

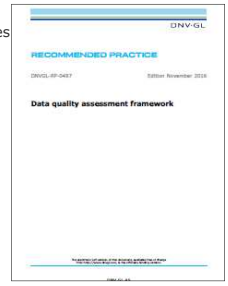
Data quality maturity levels and perspective

- **Initial**
  - Data is managed as a requirement for the implementation of projects
- **Repeatable**
  - There is awareness of the importance of managing data as a critical infrastructure asset.
- **Defined**
  - Data is treated at the organizational level as critical for successful mission performance
- **Managed**
  - Data is treated as a source of competitive advantage
- **Optimized**
  - Data is seen as critical for survival in a dynamic and competitive market

Ungraded

Why Explore the Data Quality Assessment Framework

- An opportunity for DNV GL to understand current practices, standards and processes and whether this RP need further development for this field
- An opportunity for laboratories have a fresh set of eyes with experience from other industries to assess and improve data quality management
- Assess whether this approach could provide assurance between laboratories that wish to share data?


DNV·GL

RECOMMENDED PRACTICE

DNVGL-RP-0497 Edition November 2016

Data quality assessment framework

Ungraded

www.dnvgl.com

SAFER, SMARTER, GREENER

Ungraded