Global Alliance for Genomics & Health

GA4GH Pharmacogenomics Data Standardization & Storage

Agendas & Meeting Minutes

This is the running agenda and minutes document for the GA4GH Pharmacogenomics Study Group exploring standardization and storage of pharmacogenomic data. Study groups are formed to investigate whether a new GA4GH product is needed, could be developed, and would fulfill a need not met by any existing efforts. Study group activities include broad outreach, stakeholder engagement, consultation with the Regulatory & Ethics and Security Work Streams, landscape analysis, use case development, and defining a problem statement.

For further information or to get involved, please contact info@ga4gh.org.

Table of Contents:

Meeting Protocols

- Please note that by participating in meetings, attendees agree to adhere to the GA4GH Standards for Professional Conduct.
- Meetings may be recorded for note-taking purposes. Recordings will be deleted within three months of the meeting taking place.

Relevant Articles

McDermott JH, Wright S, Sharma V, Newman WG, Payne K, Wilson P. Characterizing pharmacogenetic programs using the consolidated framework for implementation research: A structured scoping review. Front Med (Lausanne). 2022 Aug 18;9:945352. doi: 10.3389/fmed.2022.945352. PMID: 36059837; PMCID: PMC9433561.

Global Alliance for Genomics & Health

McDermott JH, Sharma V, Keen J, Newman WG, Pirmohamed M. The Implementation of Pharmacogenetics in the United Kingdom. Handb Exp Pharmacol. 2023;280:3-32. doi: 10.1007/164_2023_658. PMID: 37306816.

2024-03-26

Chairs: Videha Sharma
Attendees "Name (Affiliation)": Pedro Caraballo, Ann Moyer, Beatrice Amos, Vicky Pratt, Wes Goar, William, Mark Santcroos, Ana, Brett (Boston), Catherine Procknow, Ian Fore, Jessica Keen, Katherine Taylor, Mark Bartlett, Nicole Bertram, O Nyangiri, Robert Freimuth, Vicky Pratt, William Sproviero, Vita Rovite, Grant Wood

	Agenda Item	Person	Time
	Welcome: GA4GH Connect meeting - Please introduce yourself in the chat - Mailing list has been set up (if you are not in it or unsure please email beatrice.amos@ga4gh.org) - Study group update	Beatrice	2 mins
	Introduction and updates	Videha Sharma	10mins
	Guest speaker: Dr Ann Moyer: Ann M. Moyer, M.D., Ph.D. - Mayo Clinic Faculty Profiles	Dr Ann Moyer	20mins
	Guest speaker: Dr Pedro Caraballo: https://www.mayo.edu/research/faculty/caraballo-pedro-j-m-d/bio-20324701	Dr Pedro Caraballo	20mins

Link to meeting recording:
https://us02web.zoom.us/rec/share/vA0UbGdDWhCIrkenB7OrjqXyoTEwM0-RuY3ZCs8zY4zVUW93PodgCcoLwyNqzx7W.Os7Xkxp14mvSQxjj?startTime=1711464647000
Passcode: 92Y=d*Qy

Meeting summary:
The main focus of a study group mentioned - this meeting aims to progress on the goals and processes of the GA4GH study group, which is to identify challenges, engage a broad

cross-section of the community, and determine the need for new work as well as the effective use of existing global products if applicable.

Dr. Sharma, the PGx co-lead presented some slides to provide context for the discussion, emphasizing the importance of developing standards from an implementation perspective. He shared insights from a recent NHS England strategy launch aimed at accelerating genomic innovation, particularly in pharmacogenomics and medicines optimization. One of the key challenges identified was the role of IT infrastructure and electronic health records in supporting the implementation at scale.

The development of a standardized data model for pharmacogenomic results was discussed. This model has been reviewed internally and is set to be reviewed by the international community. The goal is to create a vendor-neutral, standardized way of storing and accessing pharmacogenomic test results across different healthcare settings. This includes developing a persistent standard for pharmacogenomic data, rather than focusing solely on messaging standards.

Questions about the intersection of this group's work with existing efforts in HL7's clinical genomic standards raised. Bob offered to provide an update on HL7's work and highlighted a separate effort he leads, which focuses on developing result-agnostic data models. These models aim to separate observational data related to patients from definitional data found in knowledge bases, facilitating a more flexible and comprehensive approach to genomic data.

The discussion underscored the need for standardized data models to support the widespread implementation of pharmacogenomics in healthcare. There's a need for clear definitions and scope.

Dr Ann Moyer presentation:
Dr Moyer presented the intricacies and challenges of pharmacogenomics, particularly emphasizing the need for consistent laboratory practices and standardization. She highlighted how pharmacogenomics, which often uses targeted genotyping, can impact patient care by identifying actionable genetic variants that influence drug response. However, a major challenge lies in ensuring that test results are consistent across different laboratories, as historically, testing focused on single genes rather than comprehensive panels.

She explained the workflow in the laboratory, starting from testing individual genetic variants to determining star alleles and predicted phenotypes. She stressed the importance of assay design, noting that different labs might select different variants, leading to variability in results. The need for standard formatting and discrete data fields was emphasized to facilitate the use of clinical decision support tools.

The variability in how data is reported by different manufacturers and the complexities of interpreting genetic data was mentioned. Moyer noted that differences in allele nomenclature and the presence of reverse strand genes add to the complexity. Furthermore, the introduction of activity scores and the variability in reporting copy number variations, particularly for genes like CYP2D6, were highlighted as areas needing standardization.

Efforts to harmonize practices in the field, such as the use of standardized databases and guidelines from organizations like the Association for Molecular Pathology and CPIC, were acknowledged. Moyer also discussed the move towards sequencing over targeted genotyping, though sequencing also presents challenges in interpreting rare variants without a standardized system.

She shared a peek at future directions in pharmacogenomics, emphasizing the need for continued standardization, better reporting practices, and integration of pharmacogenomic data into clinical decision-making processes to improve patient care outcomes.

Dr Pedro Caraballo presentation:
Dr. Pedro Caraballo's presentation focused on the integration of pharmacogenomics into clinical practice, emphasizing its importance in enhancing drug efficacy and minimizing adverse effects. He highlighted the complexity of this integration, noting that clinicians must consider various factors such as renal and liver function, drug interactions, and patient history. The relevance of pharmacogenomics is growing, particularly in cancer treatment, with a significant number of new drugs being developed based on genetic information.

Successful implementation requires a multidisciplinary approach, involving centralized governance that integrates education, clinical approval, and knowledge management. Dr. Caraballo emphasized the need for pharmacogenomic data to be integrated into electronic health records (EHR) and supported by clinical decision support systems (CDS). This ensures that clinicians receive relevant alerts and recommendations at the point of care. At Mayo Clinic, the transition to using the Epic system consolidated their knowledge and provided additional functionalities. This system includes translation tables and a genomic indicator, helping clinicians understand and apply genetic data in their practice.

Challenges remain, particularly the debate on whether academia or commercial entities should lead in standardizing and implementing these systems. Dr. Caraballo stressed the need for practical solutions to manage the extensive data involved, even if these solutions are initially imperfect. Clinicians often face alert fatigue, making it crucial to tailor alerts effectively and provide concise, relevant information with access to detailed reports when needed.

A survey revealed that while 70% of providers found pharmacogenomic information potentially useful, opinions were divided on its positive or negative impact. Dr. Caraballo concluded by highlighting ongoing efforts to improve these systems.

Global Alliance for Genomics & Health

2024-01-30

Chairs: Videha Sharma

Attendees "Name (Affiliation)": Beatrice Amos, Eleanor Lewis, Ian McNicoll, Bob Vaughan, Emma Magavern, Grant Wood, Greg taylor, Jen Harrow, Mark Santcroos, Omar Khan, Mike Cariaso, Surakameth

Agenda Item	Person	Time
Welcome, Housekeeping: - GA4GH Connect meeting - Please introduce yourself in the chat - Mailing list has been set up (if you are not in it or unsure please email beatrice.amos@ga4gh.org) - Call for presentations about current challenges from the community - to present at next meeting	Beatrice	2 min
Introduction and updates	Dr Videha Sharma	20mins
OpenEHR demo	Dr Ian McNicoll	10mins
A.O.B.	All	5 min

Link to meeting recording:
https://us02web.zoom.us/rec/share/WcxRXdu2wqzvDT9wEOlk64at_q7a1Pelm04Y7kyoMFFjGudw8LucCz4VnLJQCG74.vlxxjidPChFPjJPi

Meeting summary:
Feedback on the pharmacogenomics standards development topic is positive, with a recognized global need for a standard. Videha discussed plans for funding and outlined tasks for study group process, including a landscape analysis and engagement with potential adopters.
Will be working on use cases and potential adopters, aiming to present the landscape analysis to the product review committee in April.

The focus was on developing something fit for use in electronic health records and clinical practice. There was a debate about the HL7 FHIR standard, with a suggestion that it is suitable for data transmission but not storage. Videha talks about the separation of genetic test results and therapeutic implications in the data model. He outlined their approach of linking genotype data to therapeutic implications using an open API.

Some questions raised about standardization and the responsibility for storing therapeutic implications. The HL7 standard was acknowledged as crucial for data transmission, but not clear whether it should be used for long-term data storage in patient records.

Brief discussion about the current status of Electronic Health Record (EHR) systems in terms of their ability to store phenotype information. Grant raised concerns about EHR systems potentially being a bottleneck due to their lack of readiness for storing phenotype data. Videha shared insights from his work, proposing the integration of third-party clinical decision support providers using CDS hooks and the FHIR standard to implement in-workflow pharmacogenetic alerts within EHRs. The use of APIs for presenting data to clinicians during the prescribing workflow was emphasized, aligning with the broader trend in health IT towards API-based data presentation rather than data movement between systems. The importance of standards and agreement on data representation and messaging within the healthcare domain was highlighted as a key focus of this group.

Ian, a former GP and current data modeler discussed the challenge of integrating pharmacogenomic (PGX) codes into existing Electronic Health Record (EHR) systems. He emphasized that simple coding, even with established systems like SNOMED, might not be sufficient, as it requires potential changes to the underlying data structures of EHRs. He proposed the concept of OpenEHR, suggesting a separate data store alongside traditional EHRs to handle novel data types efficiently. This approach involves using APIs, including FHIR and OpenEHR, to access data through applications while recognizing the limitations of existing EHR systems.

Ian shares his experience in building data models for PGX and other domains using OpenEHR technology. He highlights the importance of simplifications for clinical use, focusing on key actionable information like genes, diplomas, and metabolizer status. He introduced the tool used for building data models and emphasized the non-engineering nature of the task, making it accessible to individuals with clinical backgrounds. He demos the review process for these data models, involving a tool called Clinical Knowledge Manager, which supports version control, governance, and collaborative reviews within the OpenEHR community.

Ian explained the process of reviewing data models within the OpenEHR framework. The review involves assessing metadata and data points, providing comments, and making recommendations. The participants are encouraged to focus on pertinent aspects and not feel compelled to comment on every section. The review tool, Clinical Knowledge Manager (CKM), is introduced as a platform for collaborative reviews with version control and governance.

Grant asked about representing phenotypes in OpenEHR, handling cases where metabolizer status is determined by multiple genes, and the use of identifiers under the hood. Ian addressed these questions, explaining how OpenEHR supports multiple representations of gene names, handling phenotype complexities, and addressing the challenges posed by combinatorial explosion in representing metabolizer status in SNOMED. The use of internal codes and the

ability to include structured files (such as VRS data) in the model were also discussed. Ian acknowledged the need to explore examples of VRS integration in future reviews.

2023-11-22

Chairs: Videha Sharma, John McDermott
Attendees "Name (Affiliation)": Beatrice Amos (GA4GH), Ian McNicoll (CEO, freshEHR Clinical Informatics Ltd., Director, openEHR Foundation), Rebekah Butterfield, MPH (Innovation Scientist, Oracle Health Innovation and Scientific Advisory), Mark Santcroos (Leiden University Medical Center) , Michele Mattioni Velsera), Monica Munoz-Torres, Katherine Taylor, Nicole Bertram (Epic), Angela Page, Ariel Pradipta, Brett Johnson, Eleanor Lewis, Emma Magavern, Grant Wood, Irene Muchada, Michele Mattioni, Mogomotsi Matshaba, Mrinal Thomas, Olha Nikolaieva, Omar Khan, Oscar Khan, Sean MacBride-Stewart, Sophie Harding, Stefan Nicolet, Terah Collins, Wes Goar, Xuelu (Jeff) Liu, Emily Lyga, Irafi

ACTIONS: Everyone: Please email Beatrice Amos if you have current challenges or implementations you'd like to share with the group. Or if there is an individual or group that you would like included in this meeting.

	Agenda Item	Person	Time
	Welcome	Beatrice	2mins
	Introduction and updates	Co-leads: Videha Sharma, John McDermott	20mins
	Use case presentations: 1) Supporting PGx CDS with OpenEHR 2) Oracle Health and Life Sciences 3) mijnDNAmedicatiepas.nl	Dr Ian McNicoll Rebekah Butterfield, MPH Mark Santcroos	10mins 10mins 10mins
	Open discussion and next steps		10 mins
	AOB		

Link to meeting recording:
https://us02web.zoom.us/rec/share/79mWkwqdWpqLtemN9dnKvD5s-BVbQilIabID28P4GhKmjGGzrYfBjpnVwltbEe4q._qkJTk1LE_rkLcGa

Minutes:

Dr. Ian McNicoll presented the model's framework, discussing how it would accommodate genetic information and the plan for archiving detailed variant data. There was an emphasis on reviewing and updating the model using a Delphi-like tool, with an intention to capture tested alleles and trigger updates for evolving guidelines.
Question about representing partial test results and updating clinicians about changes in guidelines. There is a need for a balance between the complexity of the model and its practical implementation.

Rebekah Oracle Health discussed her journey into multi omics, particularly pharmacogenomics, and the challenges of integrating diverse health data for clinical insights. She emphasized the need to bridge gaps in pharmacogenomics implementation across various stakeholders—patients, clinicians, laboratories, and researchers. The focus is on making complex genetic data accessible and actionable, ensuring ease of interpretation for patients and clinicians alike. She highlighted the complexities in data storage, ontology, and standardization, aiming for an EHR-agnostic, cloud-flexible system to combine laboratory insights, clinical interfaces, and patient understanding. Also the importance of evidence generation for reimbursement and equity purposes and envisions a comprehensive precision omics approach aligning various data types and standards for meaningful insights.

Mark is working on a project related to genetic variants, particularly in pharmacogenetics. Mark discussed a unique "transport mechanism" involving a QR code-embedded genetic profile on cards distributed to patients. These cards provide information on actionable variants in pharmacogenetic genes, linking to a webpage with detailed advice aligned with working group recommendations.
The focus was on decentralized data storage, with the QR code containing genetic information, ensuring patient accessibility and the latest advice translations. Mark highlighted the challenges of data transfer between systems and the lack of standardized formats. Need for smoother data exchange standards to enable seamless information transfer across healthcare systems without excessive focus on storage details.

Questions revolved around the integration of this system into electronic health records (EHRs) and pharmacy systems, with Mark explaining varying approaches among different hospitals, including storing PDFs or directly integrating genetic information into medication safety alerts. Emphasis is on a decentralized, patient-accessible genetic information system through QR-coded cards, aiming for effective integration into existing healthcare systems while underscoring the need for streamlined data exchange standards.

Global Alliance
for Genomics & Health

2023-06-05

Chairs: Videha Sharma, John McDermott
Attendees "Name (Affiliation)": Beatrice Amos (GA4GH), Wes Goar (Nationwide Childrens Hospital), Alex Wagner (Nationwide Children's Hospital), Eleanor Lewis (BC Genome Science Centre), Mike Cariaso (formerly of SNPedia & Promethease), Steve Laurie (CNAG, Barcelona), Ariel Pradipta (Genomik Solidaritas Indonesia), Greg Taylor (Canada's Michael Smith Genome Sciences Centre at BC Cancer), Grant Wood, Jen Harrow (moved from ELIXIR to AstraZeneca recently and am Director of Phenomics), Jessica Nelson (Canada's Michael Smith Genome Sciences Centre at BC Cancer), Mark Bartlett (CEO of StoreGene), Brett Johnson (Global Genes), Mark Santcroos (Leiden University Medical Center), Oscar Nyangiri (Makerere University, TrypanoGEN+/H3Africa consortium), Rebekah Butterfield (Oracle Health Innovation),

Link to meeting recording

ACTIONS: Everyone: Please email Beatrice Amos if you have current challenges or implementations you'd like to share with the group.

	Agenda Item	Person	Time
1.0	Welcome, Housekeeping: - Purpose of this studygroup - What is a study group? Process review: https://github.com/ga4gh/product-process/blob/main/product-process.md - Please introduce yourself in the chat	Beatrice	5 min
2.0	Recap and set goals for today's meeting	Videha & John	5 min
2.0	Review of survey feedback	Videha & John	5 min
3,0	Current implementations	Volunteers / open discussion	30 min
5.0	Presentation on VRS in pharmacogenetics	Wes Goar	10 min
6.0	Next steps and close: - Please register for plenary in September:	Beatrice	5 min

	https://broadinstitute.swoogo.com/ga4gh11thplenary - Mailing list has been set up (if you are not in it or unsure please email beatrice.amos@ga4gh.org) - Meeting timing/cadence		

Minutes:

VS - Introduction and recap

JM - Review of survey feedback to gain understanding of data standard use in pharmacogenetics: ~50% of survey respondents were implementing pharmacogenetics in their institutes. Lots of variability in how labs use and store data - no consistent use, except use of VRS.

VS/JM/BA - thank you for participating in survey - we may reach out to some participants to get a better understanding of use case and experience.

JM - examined what type, extent and granularity of data that needs to be captured. Highlights difficulty and challenges posed by capturing the pharmacogenetic result is variability of genes in nomenclature. Need to develop a general model to capture breadth of data.

GT - questions whether features larger than genes would be captured?

JM - not captured here, as aim is more about capturing germline pharmacogenetic variation. Should scope be extended?

GW - Suggests inviting representatives from labs to share experience. Also experts from FIHR, where similar work has already been done, and what this group could add.

VS - FIHR model links drug to gene result, which could be seen as different from a pharmacogenetic test result. These may change over time, and representing these as a pharmacogenetic profile for informing prescribing, may be something that will need to be added to the conversation.

RB - tech is looking at ways to support this work in the health/omics field.

JM - demonstrates CYP2C19 result example. Discussion of how much granularity to encode?

AW - what is a reference diplotype in slide, as reference genomes are haplotypes.

JM - may need to be amended, based on VRS presentation. Section was created for demo purposes.

SL - highlights value of capturing all or as many cases as possible. Do we need to think about gene-gene interactions within pathways?

JM - implementation currently single gene-drug interactions.

SL - what about drug evidence testing on males/females and ethnicity biases? GW has not experienced any biases in testing in clinical settings towards males.

MS - value and need of allele covering. would like to see more of what was done in HL-7. Would be good to start collecting resources.

MB - highlights challenges with confidence around variant calls / CNVs when analyzing whole genome sequence data, and quality and coverage of data being output from genome sequence.

How confident of a variant presented from tools - suggests a confidence score. Has model been stress-tested? JM - not yet.
GT - allele variant -do we have preferred transcript for this?
JM - not included yet, but necessary to include.
MC - abstraction layer required when thinking about platforms.
SL - need to explore overlaps with metadata study group.
WG - VRS in Pharmacogenetics presentation - can make slides available
https://github.com/ga4gh/vrs-python/blob/main/notebooks/PGx.ipynb
MB - will share models from survey

2023-04-13

Chairs: Videha Sharma, John McDermott
Attendees "Name (Affiliation)": Lindsay Smith, Melanie Courtot (OICR), Wes Goar (Nationwide Childrens Hospital), Omar Khan (NHS England), Nicole Bertram (Epic), Katherine Taylor (Epic), Mark Santcroos (Leiden University Medical Center), Alex Wagner (Nationwide Children's Hospital), Rebekah Butterfield (Oracle Health Innovation), Michele Mattioni (Velsera), Oscar Nyangiri (Makerere University, TrypanoGEN+/H3Africa consortium), Benjamin Haibe-Kains (UHN/PMH), Adriana Malheiro, Andy Yates, Grant Wood, Jen Harrow, Ian McNicoll (freshEHR), Bob Freimuth (Mayo), Liguo Wang, Emma Magavern.

Meeting Recording

Materials: Slides

ACTIONS:
- Everyone: Please email lindsay.smith@ga4gh.org if you have current challenges or implementations you'd like to share with the group
    - Wes Goar & Alex Wagner: will share work modeling PGx star alleles and diplotypes using VRS during. Can demonstrate how we can accurately represent what is precisely known from the genetic/clinical testing and how that maps to PGx star alleles.
    - Adriana: presentation on how MedGen, GTR and MGS can be leveraged in PGx implementation:
        - MedGen can be used to standardize phenotype terms and facilitate data sharing across all platforms (EHRs, labs, patient resources)
        - GTR can provide access to available PGx tests and can support standardization of descriptions of PGx tests
        - Medical Genetics Summaries for access to actionable PGx information tailored to clinicians including therapeutic guidelines based on genotype from authoritative sources like FDA drug labels, CPIC, DPWG, CPNDS and others
    - Mark Santcroos: dutch initiative on a DNA medication card (https://mijndnamedicatiepas.nl)
    - Eric Roller: how to represent the variation for star alleles when there are gene

duplications on one haplotype. CYP2D6 is the example I had in mind where duplication is a common variation in the population.
- Lindsay: To send summary email and schedule a call in 4 weeks
- John & Videha: To prepare and circulate some initials thoughts for feedback

	Agenda Item	Person	Time
1.0	Welcome, Housekeeping - GA4GH Connect is next week! - Call for GA4GH Driver Projects - GA4GH Study Groups - Please introduce yourself in the chat	Lindsay Smith	5 min
2.0	Setting the stage for GA4GH Pharmacogenomics Data Standardization [Presentation]	Videha Sharma, John McDermott	10 min
3.0	Challenges in pharmacogenomics data standardization and storage	Group Discussion (led by Videha)	30 min
4.0	Call for presentations about current challenges from the community - to present at next meeting	Videha & John	5 min
5.0	Meeting logistics moving forward - Mailing list has been set up (if you are not in it or unsure please email lindsay.smith@ga4gh.org) - Meeting timing/cadence	Lindsay	5 min
6.0	A.O.B.	All	5 min

Minutes:

Setting the scene: development of data standards for pharmacogenetics
JM: Poor medicines effectiveness and adverse reactions are a problem, and factors that contribute to this are poor adherence and predisposing pharmacogenomics variation. 99% of individuals carry pharmacogenomic variance.
Despite evidence for multiple gene-drug pairs, there is a translational gap (particularly in the NHS). In the UK, only 4 of 26 gene-drug pairs have made it into clinical practice. One of the gaps is strategy. Recognized that preemptive storage approach is a much more effective way to use that data. Taken a range of different approaches, but they all show heterogeneity in the way tests are ordered, used, and the return of results - which makes sense in the context of diverse healthcare services.

This type of genetic data is different in the ways its used compared to other genetic data. The particular challenge that many healthcare organizations face, is how to move that data around a disparate healthcare service. The data is in a lab, but how to move that around. There might be secondary care using EPIC, or CERNER, or EMIS. How to create a process to ensure in which data moves around that service interoperable, and make sure delivery is equitable.
We need to decide how the data is stored within data repositories that is consistent.
VS: There is a good bunch of literature out that highlights the importance of standardizing pharmacogenetic data. Specifically around health data standards, where only 4/32 studies actually talked about data standards as part of their implementation process.
Be interoperable by design - agree on a common language. For PGx specifically, evidence that lack of interoperability as a barrier to implementation at scale. PGx is looking to provide clinical decision support, so important to store data in a structured way. The most appropriate way would be to think about PGx data in a vendor-agnostic way, and store persistent data using a platform strategy.
Speaking to clinical stakeholders and patients, there are a few requirements: genomic data is for life, PGx data will be used across healthcare settings and across a range of purposes. Data will be valuable for research and analytics in the future. Propose using openEHR and FHIR to support design of this.
The need for this group - understand current practice and share learning, bring clinical and informatics community together, reviewing existing models and identify gaps, and develop 'pharmacogenetic test result' data model - review and feedback. First place to start is to think about a data model. Have a rough strawman idea around this.

BH: do you consider the associations between tumor somatic aberrations and drug response / adverse side effects relevant for this group? Or only germline?
JM: My feeling is that this is related to germline, but not to say that somatic is very relevant. Two very different use cases, and would treat them differently.
BH: In a way, the standards might be very similar.
JM: With regard to the nomenclature, there's activity scores and star alleles that aren't present when you're thinking about targeted therapy. Potentially need distinct groups.
NB: If thinking about therapeutic recommendations, it does open up to a number of different things that would need to be represented. I think it makes sense to separate at least for initial discussions, with the idea to expand later.
JM: Seems like a sensible approach. Start with germline complexity.
AW: Add my agreement, we spend a lot of time thinking about structuring cancer therapeutic knowledge. There's a a lot of similarities and common patterns in how we structure knowledge, and we could re-use those patterns, but there are major differences.

AM: To standardize phenotype nomenclature, have you considered using MedGen which includes SNOMEDCT (as well as UMLS, mondo, HPO, OMIM) and has drug response terms
VS: Thanks for raising, haven't come across that yet. Will take away and look at that.
AM: We use MedGen, have to create phenotypes for drug responses. Integrate all of the sources together for phenotypes, then if there is no source, we create our own, give them an

identifier. For MedGen, we try to put everything together. For SNOMED, we find there weren't good terms to identify drug responses.
VS: Is MedGen a standard or a platform?
AM: We bring the sources together, only create terms for which there aren't any. Harmonize what the community uses.
Example: https://ncbi.nlm.nih.gov/medgen/382487
JM: Really useful resource, and potentially an area to pull data from. An additional archetype is needed to store this data.
AM: Have links from MedGen that go to ClinVar, for example.
IM: This is hugely valuable background information. The gap we think we'd identified, was identifying the enzyme status as a phenotype, not a drug. The impact on a specific therapeutic. Complication that a single PGx test is not necessarily all encompassing, might only test a subset. Emerging phenotype as people get more and more tests done.
JM: Finding a way to adequately record the PGx testing done. Approach we've taken for now, is record the test, but something needs to be done around that challenge.
VS: Keen to explore that balance between being perfect and being pragmatic.
MS: We generally approach this with something called coverage, what was the test able to see and then combine with the variants. Agnostic of how the test was done. Wanted to comment on the persistence. There are different opinions of what persistence is, arguably the only thing that is really persistent is genetic makeup. Star alleles change over time. Want to make the translation as late as possible. See a hierarchical model where you add or amend information.
JM: Spot on, the only thing that is true is the raw genetic data. Storing your results in an EHR or anywhere, and then trying to use that same result in 20-30 years time may not work.

VS: Thinking about the granularity, what needs to persist.
Open call for presentation or implementations to share/present.

JM: Put some thoughts down on paper and circulate. Then come back together.
Then plan ad hoc meeting, roughly 4 weeks.

BF [from chat]: IMO, the star allele system is better for a closed world system where all knowledge is known a priori. In reality, as Emma just indicated, this can be an issue in PGx implementations because the PGx domain is still evolving and there is a lot we don't yet know (open world). Personally, I view star alleles as a type of interpretation, which layers over primary genomic observations/results and which may change over time as new alleles are discovered.

How will this group build on/complement existing efforts in this space. Thinking specifically about CPIC and the dutch group. Curious about anyone had opinions on that, or can talk about next steps
JM: I think that's important. I'm a member of CPIC myself. They are the leaders of standardizing nomenclature. Get input from this group, and share with them our progress and their approval is key.

Global Alliance
for Genomics & Health