# 2022-05 Clinical Genomics: Retrieving FHIR Genetic data with Operations

- **Short Description**
  - Using/accessing FHIR Genetic data: this track is about what/how you can access data to meet clinical use case. Operations expand the capabilities of a FHIR server to deliver genomics functionality to answer key clinical queries in a streamlined fashion, that is agnostic to the format used to report genetic data (i.e. the input parameters use a single format modality rather than an application needing to know how the data is stored in the server). New operations are proposed that return associated Diagnostic or Therapeutic implications – answering such a request as ‘return the diagnostic implications associated with a specific genetic variation’ or ‘return the diagnostic and therapeutic implications of patients that match my patient genetically’ without needing to know the variant formatting on the server.

- **Long Description**
  - Operations expand the capabilities of a FHIR server to deliver genomics functionality. Even with conformance to the Clinical Genomics IG, genetics can require dealing with liftover, synonymous HGVS, subsumed variants, overlapping variants, mapping between system representations, and multiple formatting modalities. Application developers, or other users of genetic data, will appreciate how operations simplify development effort, providing a single input format and streamline access. This track is about using FHIR genomics data. Through understanding the gaps in usability we will build a better more usable standard.

- **Type**
  - Test a FHIR-associated specification

- **Submitting Work Group/Project/Accelerator/Affiliate/Implementer Group**
  - This track will be hosted by CodeX’s GenomeX FHIR Accelerator project (Operations Use Case) in collaboration with the Clinical Genomics Work Group with input from any interested parties.

- **Track Lead(s)**
  - Bob Dolin BDolin@elimu.io
  - Bret Heale bheale@humanizedhealthconsulting.com

- **Track Lead Email(s)**
  - BDolin@elimu.io, bheale@humanizedhealthconsulting.com

- **Related Tracks**

- **FHIR Version**
  - Current Build, FHIR R4
Specification of focus

- The operations build on the Clinical Genomics Reporting IG https://build.fhir.org/ig/HL7/genomics-reporting/branches/operations/operations.html
- Note: Operations return data as FHIR profiles compliant to the Clinical Genomics Reporting IG.
- Reference implementation server for FHIR Genomics Operations found at: https://fhir-genomics-apis.herokuapp.com/
- Description of Reference Implementation (patient lists, capabilities, etc... of reference server): https://docs.google.com/document/d/1b6u-AfjPKh0lZ33nLCT7jUMUfFHuh3Sk5OuDyVg/edit#
- Terminology Servers this track uses? None
- Specific Terminology Code Systems and Value Sets needed for this track? Specified in IG
  - Value Sets: https://build.fhir.org/ig/HL7/genomics-reporting/artifacts.html#terminology-value-sets

Artifacts of focus

https://build.fhir.org/ig/HL7/genomics-reporting/branches/operations/operations.html
Reference Implementation: https://fhir-genomics-apis.herokuapp.com/
Description of Reference Implementation (patient lists, capabilities, etc... of reference server): https://docs.google.com/document/d/1b6u-AfjPKh0lZ33nLCT7jUMUfFHuh3Sk5OuDyVg/edit#

Expected participants

Application developers, EHR vendors, Business and Data Analysts/Scientists

Zulip stream

https://chat.fhir.org/#narrow/stream/179197-genomics/topic/Connectathon.2030.20.-.20Genomic.20operations

Track Kick Off Call

https://us02web.zoom.us/j/85646519661
04/28/2022
10 am - noon Eastern

Track Kick-off and Orientation - brief overview on operations and orientation to Track, scenarios and CG IG Discuss resources available Solicit additional scenarios - this is your connectathon, what would you like to test out that’s not covered? Follow-up discussions on Zulip thread or confluence.

Agenda

<table>
<thead>
<tr>
<th>Date</th>
<th>Start Time [Eastern]</th>
<th>End Start Time [Eastern]</th>
<th>Session Title</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>05/03</td>
<td></td>
<td>Build/Test/Validate; Office Hours</td>
<td>Connection day 1 - see description and track for details</td>
<td>One meeting/all, on all day. 9:00 am - 10:00 am Eastern Time Zulip channel is a must! Please feel free to use it. 10:00 am - 10:15 am EST (educational/breakout) Focus on scenario 0 - why operations 10:15 am - 10:45 am EST (educational/breakout) Focus on implcations operations - why operations 11:00 am - 1:00 pm EST Testing scenarios open office hours/work sessions/troubleshooting - issue handling 1:00 pm - 2:00 pm EST Bring your own topic 2:00 pm - 3:00 pm EST Testing scenarios open office hours/work sessions/troubleshooting - issue handling 3:00 pm - 4:30 pm EST Demo or descriptions of days efforts - regroup of participants 4:30 pm - 5:30 pm EST Break 5:30 pm - 7:30 pm EST Late zoom session When zoom closes, the zulip channel will still be live - please feel free to use it.</td>
</tr>
<tr>
<td>05/04</td>
<td></td>
<td>Build/Test/Validate; Office Hours</td>
<td>Connection Day 2 - see description and track for details</td>
<td>One meeting/all, on all day. 9:00 am - 9:30 pm EST Check in for anyone with questions, comments, or feedback 9:30 am - 11 am EST Recap day 1. Challenges/what-learned 11:00 am - 12:30 pm EST Continue scenario testing/Open office hours/work session 12:30 pm - 2 pm EST Special topics on divided by group 1:30 pm - 3:30 pm EST Task/Last Dime/Discussion Future Connections  Track report-out After conclusion Zulip channel will still be available (as are all the Clinical Genomics Zulip channels and CG calls). Join the accelerator!</td>
</tr>
</tbody>
</table>

Track Details

System roles:
The following is a call-out to the specific operations which will be included in the track.

Clients should familiarize themselves with the operations (how to use them and what to expect as a return). Clients are asked to estimate or track their development time. Bonus points for a nice visualization of the parsed and received data. If there is a specific operation that you would like to use for a specific use case or try out, please check with the track leads and they will ensure that relevant data is in the reference implementation. You can send data for inclusion in multiple formats including FHIR or VCF.

Servers should review the operations and determine which operations they will support. The track leads are available for consult and the reference implementation has an accompanying GIT hub. Please have your server prepared 2 weeks before the connectathon. Using the zulip chat you can connect with client participants and track leads maybe able to help with populating data to meet test use case.
Subject Operations | Population Operations

**Genotype Operations**
- simple variants: find-subject-variants; find-subject-specific-variants
- structural variants: find-subject-structural-intersecting-variants; find-subject-structural-subsuming-variants
- haplotype/genotypes: find-subject-haplotypes; find-subject-specific-haplotypes

**Phenotype Operations**
- therapeutic implications: find-subject-tx-implications
- diagnostic implications: find-subject-dx-implications

**Metadata Operations**
- study metadata: find-study-metadata

<table>
<thead>
<tr>
<th>Operation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>find-subject-variants</td>
<td>Determine if simple variants are present that overlap range(s).</td>
</tr>
<tr>
<td>find-subject-specific-variants</td>
<td>Determine if specified simple variants are present.</td>
</tr>
<tr>
<td>find-subject-structural-intersecting-variants</td>
<td>Determine if structural variants are present that overlap range(s).</td>
</tr>
<tr>
<td>find-subject-structural-subsuming-variants</td>
<td>Determine if structural variants are present that fully subsume a range.</td>
</tr>
<tr>
<td>find-subject-haplotypes</td>
<td>Retrieve haplotypes/genotypes for specified genes.</td>
</tr>
<tr>
<td>find-subject-specific-haplotypes</td>
<td>See if specified haplotypes/genotypes are present.</td>
</tr>
<tr>
<td>find-subject-tx-implications</td>
<td>Retrieves genetic therapeutic implications for variants/haplotypes/genotypes.</td>
</tr>
<tr>
<td>find-subject-dx-implications</td>
<td>Retrieves genetic diagnostic implications for variants.</td>
</tr>
<tr>
<td>find-population-specific-variants</td>
<td>Retrieve count or list of patients having specified variants.</td>
</tr>
<tr>
<td>find-population-structural-intersecting-variants</td>
<td>Retrieve count or list of patients having structural intersecting variants in specified regions.</td>
</tr>
<tr>
<td>find-population-structural-subsuming-variants</td>
<td>Retrieve count or list of patients having structural subsuming variants in specified regions.</td>
</tr>
</tbody>
</table>

**Scenarios**

These scenarios are examples of using the operation to perform tasks. In the last scenario, we provide some additional examples of use cases for operations. The goal is to test out the operations and gain familiarity with them. Please feel free to work on any additional scenarios that make use of the operations. We hope to gain feedback on utility, completeness and ways to improve or extend the operations!

**Security and Privacy Considerations:**

No security or privacy considerations for use of the reference server. The data is synthetic. If implementing a server for Genomics Operations be sure to use test patient data and not expose real patient information.

**Scenario #0 Name “Variant Normalization - encapsulation through operations”**

Scenario brief description

"Patient has a variant in APC gene (see [https://www.ncbi.nlm.nih.gov/clinvar/variation/819274](https://www.ncbi.nlm.nih.gov/clinvar/variation/819274)). Search for this variant in patient m123, using the find-subject-specific-variants operation. Modify your search to use any of the variant synonyms on the Clinvar page, to see that the variant is still identified and returned."

Variants can overlap or be subsumed by other variants. Also, optionally in the Clinical Genomics FHIR Variant Profile allows an implementor to represent variants in different ways. As an example the following are synonymous:

- component:dna-chg: HGVS = NM_001195798.2:c.12G>A
- component:genomic-dna-chg: HGVS = NC_000019.9:g.11200236G>A
- component:genomic-dna-chg: HGVS = NC_000019.10:g.11089560G>A

Using Multiple components (VCF-like):

- component:genomic-ref-seq: NC_000019.10
- component:ref-allele: G
- component:alt-allele: A
- component:coordinate-system: 0-based interval counting
- component:exact-start-end: start = 11089559

To be complete, a FHIR query could need to comprehensively include all the different HGVS for a variant, the VCF-like use of multiple components, and perhaps even a list of variants that contain the variant of interest as input parameters. The operations used here in this Track all use a single format for input parameters and guarantee that the search will cover all the possible HGVS and other formatting options. Implicit in these FHIR Genomics operations is that variants in requested regions are returned regardless of how they are formatted/represented stored in a server. Please read [https://build.fhir.org/ig/HL7/genomics-reporting/branches/operations/operations.html#expectation-of-normalized-variant-search](https://build.fhir.org/ig/HL7/genomics-reporting/branches/operations/operations.html#expectation-of-normalized-variant-search) for more information. "Liftover" is also performed by the server providing the operations.

Action (Client):

Step 1. Perform find-subject-specific-variants operation to retrieve data.
Step 2. Perform the operation again, using a different synonym of the variant. Repeat using several synonyms.

Action (Server):
Step 1. Populate server with data.
Step 2. Implement and provide access to the operation.

Precondition: Success Criteria:
Operations functional and data is in server. Client is prepared to send REST call and receive data. (end-user UI visualization gets you bonus points).

Scenario #1 Name "Recent knowledge base update"
Scenario brief description
"Over the course of the past two years, many patients have been tested for variants pathogenic for hereditary breast and ovarian cancer syndrome (HBOC), Lynch syndrome (LS), and familial hypercholesterolemia (FH). Recent knowledge base updates have added to the list of known pathogenic variants, and have reclassified the pathogenicity of many variants. We now want to find patients that have pathogenic variants for HBOC, LS, or FH, based on the new knowledge base."

Action (Client):
Step 1. Perform operation to retrieve data.

Action (Server):
Step 1. Populate server with data.
Step 2. Implement and provide access to the operation find-population-dx-implications

Success Criteria:
Data is retrievable and retrieved.
Please share operation used, screenshot of result, development time, time for transaction. (Client)
Please share server trace log of the transaction and transaction time. (Server)

Bonus point:
Share screen shot of visualization of result. Comments and critique of the operation.

TestScript(s): N/A

Scenario #2 Name "Population find Cohort matching clinical trial study criteria in genomics"
Scenario brief description
"A researcher has developed a new drug, designed for cancer patients with large deletions involving all or part of BRCA1 (NC_000017.11:43044294-43125364) or BRCA2 (NC_000013.11: 32315507-32400268), and wants a list of potential clinical trial participants."

Action (Client):
Step 1. Perform operation to retrieve data.

Action (Server):
Step 1. Populate server with data.
Step 2. Implement and provide access to the operation find-population-structural-intersecting-variants

Success Criteria:
Data is retrievable and retrieved.
Please share operation used, screenshot of result, development time, time for transaction. (Client)
Please share server trace log of the transaction and transaction time. (Server)

Bonus point:
Share screen shot of visualization of result. Comments and critique of the operation.

TestScript(s): N/A
Scenario #3 Name "HLA matching and implications"

"A patient has undergone HLA typing, and clinicians wants to know if there is any haplotype or genotype indication that HLA-A*23 is present".

Action (Client):
Step 1. Perform find-subject-specific-haplotypes operation to retrieve data.

Action (Server):
Step 1. Populate server with data.
Step 2. Implement and provide access to the operation

Precondition: Success Criteria:
Operations functional and data is in server. Client is prepared to send REST call and receive data. (end-user UI visualization gets you bonus points).

Success Criteria:
Data is retrievable and retrieved.
Please share operation used, screenshot of result, development time, time for transaction. (Client)
Please share server trace log of the transaction and transaction time. (Server)

Bonus point:
Share screenshot of visualization of result. Comments and critique of the operation.

TextScript(s): N/A

Scenario #4 Name "Genetic related diagnosis of patients like mine"

"Physicians have wondered 'What do patients with my patient's variants have in common with regards to reported genetics-linked diagnostic implications?' In this scenario, use a set of variants to represent your physician's patient's variants to find a population in the server that has those variants (or various subsets). Then, use that population and find the implications associated with the variants found in the population. Try a smaller list of variants first. Be creative."

Action (Client): *suggested steps, be creative and use the operations and process that make the most sense to you
Step 1. Perform find-population-specific-variants, and/or find-population-structural-intersecting-variants, and/or find-population-structural-subsuming-variants operation, as appropriate, to retrieve list of patients.
Step 2. Iteratively perform find-subject-dx-implications.
Step 3. Compare the diagnostic implications of the population to the patient.

Action (Server):
Step 1. Populate server with data.
Step 2. Implement and provide access to the operations to meet the use case.

Precondition: Success Criteria:
Operations functional and data is in server. Client is prepared to send REST call and receive data. (end-user UI visualization gets you bonus points).

Success Criteria:
Data is retrievable and retrieved.
Please share operation used, screenshot of result, development time, time for transaction. (Client)
Please share server trace log of the transaction and transaction time. (Server)
**Scenario #5 Name: "Additional Examples"**

Feel free to devise additional scenarios to test. Here are some explicit examples:

- Patient HG00403 has WES of a tumor biopsy specimen. See if patient HG00403 had any simple somatic variants detected in BRAF (NC_000007.14:140713327-140924929) or ERBB2 (NC_000017.11:139688093-39726660) by test.

- Patient HG00403 had any of these [NM_00354609.2:c.1797T>A, NM_007294.4:c.5559C>A, NM_000001.10:c.1222538T>A, NM_000038.6:c.5559G>T, NM_00354609.2:c.1802G>T] simple variants detected by test? In test2.

- Patient HG00403 has a structural variant that intersects with EGFR exon 19 or partial deletions. See if patient HG00403 has a structural variant that subsumes MET (NC_000057.14:116672195-11698386) in specimen.

- Patient HG00403 underwent preemptive pharmacogenomic testing. Clinician now wants to prescribe amphotericin, but first wants to see CYP2D6 and CYP2C19 star alleles.

- A researcher has developed a new drug, designed for cancer patients with large deletions involving all or part of BRCA1 (NC_000010.11:116672195-140924929) and BRCA2 (NC_000013.11:23315507-3204208) and wants a list of potential clinical trial participants.

- A transplant specialist is looking for potential unrelated donors in the donor registry that are a perfect HLA match for their patient.

- A pharmacogeneticist is studying the accuracy of CYP2D6 (whole gene deletion) calling, and wants to compare cases where a structural variant caller indicates a whole gene deletion of CYP2D6 (NC_000022.10:42522500-42526812) against a pipeline used to report pharmacogenes. To begin the process, the pharmacogeneticist identifies all patients that have structural variants subsuming CYP2D6.

- A health plan establishes a quality assurance program aimed at ensuring that patients with NSCLC who are candidates for molecularly guided treatment are properly identified.

**Action (Client):**

Step 1. Perform operations to retrieve data and any additional logic for the use case.

Step 2. Populate server with data.

Step 3. Implement and provide access to the operation.

**Precondition: Success Criteria:**

Operations functional and data is in server. Client is prepared to send REST call and receive data. (end-user UI visualization gets you bonus points).

**Success Criteria:**

Data is retrievable and retrieved.

Please share operation used, screenshot of result, development time, time for transaction. (Client)

Please share server trace log of the transaction and transaction time. (Server)

**Bonus point:**

Share screen shot of visualization of result. Comments and critique of the operation.

TestScript(s): N/A
<table>
<thead>
<tr>
<th>Security and Privacy Considerations:</th>
</tr>
</thead>
<tbody>
<tr>
<td>No security or privacy considerations for use of the reference server. The data is synthetic. If implementing a server for Genomics Operations be sure to use test patient data and not expose real patient information.</td>
</tr>
</tbody>
</table>