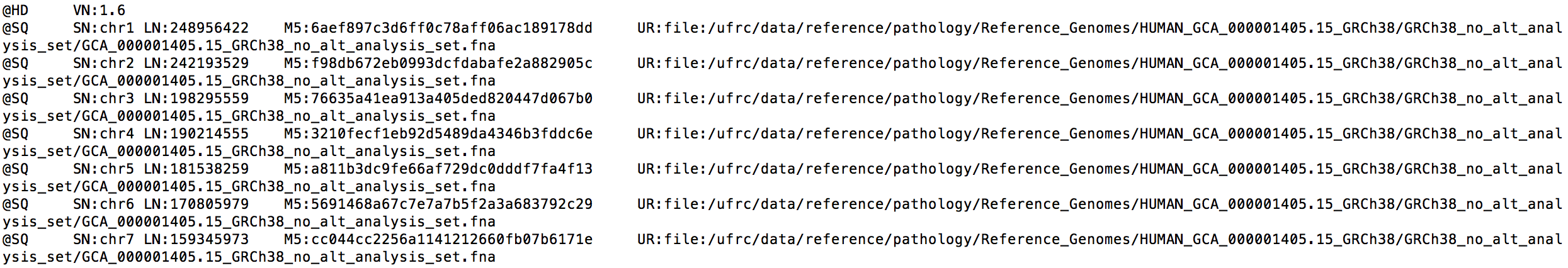
**VCF2FHIR Design Document**

(Version 1)

**Inputs:**

1. VCF file
2. Genome reference build -- HG38, HG19, ..
3. Reference Dictionary -- reference.dict file generated by Picard tool



1. Include Regions File – Bed File

* <https://en.wikipedia.org/wiki/BED_(file_format)#:~:text=The%20BED%20(Browser%20Extensible%20Data,adopted%20by%20other%20sequencing%20projects.>
* Look for python libraries – example below but you can check more.
* <https://daler.github.io/pybedtools/>

1. Non-callable regions – Bed File
2. Patient ID -- would be same as VCF file sample id; also we assume each we vcf correspond to a single sample
3. Gene Annotation File --
4. Variant types – SNVs and Indels
5. Filter Flag – filters out non-pass variants
6. Phasing

**Outputs**:

FHIR JSON file – Beow are instructions

* <http://build.fhir.org/ig/HL7/genomics-reporting/find-subject-variants.html>
* <http://build.fhir.org/ig/HL7/genomics-reporting/region-studied.html>
* <http://hl7.org/fhir/observation.html>
* <http://hl7.org/fhir/uv/genomics-reporting/StructureDefinition/variant>
* http://hl7.org/fhir/uv/genomics-reporting/StructureDefinition/sequence-phase-relationship