

2016-05-09 -5PM OutlineGenomicModel_ for HL7 - Lite Genetics report

Notes

Would have two different panels for “small” variants. (Assume) one for the simple (single allele), and one for complex (multiple allele, e.g. Haplotype or compound het, etc.). This could repeat. We could collapse them into one panel with a small variation type (e.g. simple, haplotype, compound Het, etc.)

Don’t think we need an explicit node (structure) for the genotype itself, because the “whole report” describes the genotype and we already have a section for the overall report.

Have added an attribute to the Simple variant (allele) level—the allelic frequency (a number from 0 to 1) that describes the relative frequency (as a fraction) of a given allele to other alleles at the same position. (Recall there can be more than two). The content for the large (copy number) variants remains the same.

Also note, in HL7 V2.5.1 OBX-4 is used to distinguish segments with the same Observation ID in OBX-3. And to group a set of distinct IDs that repeat together. So will only need one level of dots to deal with the levels of nesting that exist in the genomics specification?

Report Section for Variables that apply to the Overall Study.

	Data Type	LOINC # Note links	R O C	Observation display Name- draft version	Card	OBX -4	OBX-5 examples	Comments
		81247-9		HL7 genetic variant reporting panel	0..*			
		81294-1		Genetic form configuration controls				Need for controlling input into form not essential LOINC codes for inclusion in the HL7 message
A	CWE	81248-7	-	Report transcript reference sequence coding system			NCBI_NM ENSEMBLE RefSeq ENST	Answer list will include a) Ensembl transcript prefix = “ENST”, b) NCBI’s RefSeq transcript, prefix = “NM_”, c) LRG transcript prefix = “LRG_” and includes “t1” to indicate transcript, or d) Local. Local will require the entry of a question that specified the OID for that set of identifiers.
B	CWE	81249-5	-	Report genomic reference sequence coding system			NCBI_G or C	Answer list will include a) Ensembl genomic prefix = “ENSG”, b) NCBI’s RefSeq genomic, prefix = “NG_” or “NC_”, c) LRG genomic ref seq, prefix = “LRG_” or d) Local reference sequence. Local will require the entry of a question that specified the OID for that set of identifiers
C	CWE	NEW	-	Variant coding system				Needed but only for controlling form behavior Choices for now will be ClinVar Variant ID, COSMIC

2016-05-09 -5PM OutlineGenomicModel_for HL7 - Lite Genetics report

D	ST	81295-8	C	OID For local transcript reference sequence				If answer to question A is local- ask for the OID
E	ST	81296-6	C	OID For local genomic reference sequence				If answer to question B is local –ask for the OID
		81306-3		Variables that apply to the overall study				
A	CNE	48018-6		HGNC Gene(s) examined	0..*		21497^ACAD9^HGNC	Use HGNC – symbol for single or multi-gene studies. At present the look up for this option assumes Human specimens.
B	CWE	36908-2	C	Gene mutations tested	0..*			At least one of B, C or D must be included-
C	NR	51959-5	C	Ranges of DNA sequence examined	0..*			Based on the Genomic reference sequence.- may repeat if the range is discontinuous
D	TX	81293-3	C	Description of ranges of DNA sequences examined	0..1			Genetic test reports only rarely include explicit numeric ranges which are often proprietary ;so the reports more often describe the regions in narrative. E.g. “all coding regions and appropriate flanking regions”. This variable is only relevant to for sequencing studies.
E	TX	53577-3	O	Reason for study	0..*			Belongs in OBR-31 included here for convenience of from capture and to emphasize it should be part of report
F	CWE	51967-8	O	Genetic disease(s) assessed	0..*		2971795010^Deficiency of isoibutyryl-coenzyme A dehydrogenase (disorder)^SCT	
G	CWE	62374-4	C	Human reference sequence assembly and build				Needed for some kinds of refseqs. (e.g Genomics NCBI ref seqs if they don’t carry the version# and for Ensemble genomic reference sequences On build should apply to all variables in the report
H	FT;E D	51969-4	P	Full narrative report (e.g. PDF, Word Document that would look like current reports)				ED = encapsulated data which can accommodate WORD, PDFs and many other format. Further, guidance on where to identify the specific type of encapsulated data will be forthcoming.
I	ST	81303-0	P	HGVS version			2.120831	Any change in the recommendations will get a new version number based on the date of the change. Both in the version list, and on the page giving details of the change, it will be clearly marked using a format like date 2012-08-31. The version of the HGVS recommendations including

2016-05-09 -5PM OutlineGenomicModel_ for HL7 - Lite Genetics report

							that change will be version 2.120831.
J	CNE	51968-6	R	Genetic analysis overall interpretation		51968-6^Positive^LN	

Report Section for Variables that define a Simple Variant (could be more than one simple variant not related to each other).

	Data Type	LOINC #	R C O	Observation display Name- draft version	card	OBX -4	OBX-5 Example values	Comments
		81250-3		Simple variants – panel (may repeat)				
A	CWE	81252-9	R	Simple variant			30880^NM_014049.4(ACAD9): c.1249C>T (p.Arg417Cys)^ClinVar-V	Relates to 48003-8 Code + full name- which should be considered for retirement_ should
				<u>Transcript specification spelled out- needed when the variant is not yet part of a public data base such as ClinVar and/or when the reporting service wants to present them as distinct observations</u>				
B	CWE	48018-6	C	Gene :			21497^ACAD9^HGNC	
C	CWE	51958-7	C	Transcript Reference Sequence ID:			NM_014049.4^^RefSeq_NM	
D	CWE	41103-3		DNA change c.HGVS			c.1249C>T^^c.HGVS	
E	CWE	48005-3	O	Amino acid change p.HGVS:			p.Arg417Cys^^p.HGVS	
F	CWE	48019-4	O	DNA sequence variation type			LA6690-7^Substitution^LN	
G	CWE	48006-1	O	Amino acid change type			LA6698-0^missense ^LN	
				<u>Genomic specification</u>				
H	CWE	48013-7	R	Genomic Reference Sequence:			NG_017064.1^^RefSeq_CG	Required if I, J or K is present
I	ST	69547-8	P	Genomic Ref allele:			C	
J	NM	81254-5	P	Genomic Allele location:			31731	
K	ST	69551-0	P	Genomic Alt allele:			T	
				<u>Other optional codes related to simple variant</u>				
L	ID	48004-6	O	dbSNP ID:			rs368949613^^dbSNP	Almost always included for pharmacogenomics studies. dbSNP ID alone is not enough. Must be accompanied by the change e.g. A>C
M	CWE	81255-2	C	COSMIC				Usually the preferred coding system for cancer study ,COSMIC has mutation IDs for simple and for structural variations –so will likely allow IDs from either of those sources to be used
N	CWE	81257-8	O	CIGAR				Used primarily for alignment in earlier stages of

2016-05-09 -5PM OutlineGenomicModel_ for HL7 - Lite Genetics report

	Allele						genetic study analysis
				<u>Other possible attributes</u>			
O	NM	81258-6	P	Allelic Frequency NFR		0.47	This reports the fraction of all of the reads at this genomic location that were represented by the reported allele. For homozygotes it will be close to one, For heterozygotes it will be close to 0.5 (unless other variants exist as can occur with mosaics and mixtures of tumor and normal cells.
P	CWE	48001-2	O	Chromosome location of variant		3q21	Will create coding system from regions collected by NCBI
Q	CNE	48002-0	P	Genomic source class		LA6683-2^Germline^LN	Preferred for each separate variant when cancer cells involved germ line variants may be seen in normal cells
				<u>Allelic state and interpretive attributes</u>			
R	CNE	53034-5		Allelic state:		LA6706-1^Heterozygous^LN	
S	CNE	53037-8		Clinical significance:		LA6668-3^Pathogenic^LN	
T	TX	81259-4		Associated phenotype:		Acyl-CoA dehydrogenase family, member 9, deficiency of	

Report Section for Complex Variants (those with multiple alleles) -- structure can repeat because there can be more than one complex variant within one study (e.g. two haplotypes).

Sources for example: <http://www.ncbi.nlm.nih.gov/clinvar/variation/16895/> <http://www.ncbi.nlm.nih.gov/gene/1565>

	Data Type	LOINC #	R C O	Observation display Name- draft version	card	OBX -4	OBX-5 Example values	Comments
		81251-1		Complex variant – panel (if reporting multiple complex variants, may repeat)				
B	CWE	81260-2		Complex variant [Identifier]			16895^NM_000106.5(CYP2D6):c.[886C>T;457G>C] – Haplotype^ClinVar	
F	CWE	81262-8		Complex variant HGVS			c.[1749A>G ; 2549delA]^HGVS	
G	CWE	81263-6		Complex variant type		1	LAXXXX-X^Haplotype^LN	
H	CWE	81259-4		Associated phenotype			688395015^Debrisoquine adverse reaction (disorder)^SCT	
I	CNE	53037-8		Clinical significance			LA6668-3^Pathogenic^LN	

2016-05-09 -5PM OutlineGenomicModel_ for HL7 - Lite Genetics report

J	CNE	53034-5	Allelic state			LA6706-1^Heterozygous^LN	
		81259-4	Simple variant within complex variant				
A	CWE	48008-7	Simple Variant:		1.1	31934^NM_000106.5(CYP2D6):c.886C>T (p.Arg296Cys)^ClinVar	
			<u>Transcript specification</u>				
C	CWE	48019-4	DNA sequence variation type			LA6690-7^Substitution^LN	SNV – single nucleotide variant
D	CWE	48018-6	Gene:			2625^CYP2D6^HGNC	
E	CWE	51958-7	Transcript Reference Sequence ID (aka NM_RefSeq):			NM_000106.5^^RefSeq	
F	CWE	41103-3	DNA change:			c.886C>T^^c.HGVS	
G	CWE	48005-3	Amino acid change:			p.Arg296Cys^^p.HGVS	
			<u>Genomic specification</u>				
H	CWE	48013-7	Genomic Reference Sequence:			NG_008376.3^^RefSeqGene	
I	ST	69547-8	Genomic Reference (Ref) allele:			C	
J	NM	81254-5	Genomic Allele location:			42127941	Genomic location: Chr22: 42127941
K	ST	69551-0	Genomic Alternate (Alt) allele:			T	
			<u>Other optional codes related to simple variation</u>				
L	CWE	48004-6	dbSNP ID:			rs16947^^dbSNP	
M	CWE	81256-0	COSMIC				
N	CWE	81257-8	CIGAR				
			<u>Other possible attributes</u>				
O	NM	81258-6	Allelic Frequency NFR			0.40045	
P	CWE	48001-2	Chromosome region			22q13.2	
		81259-4	Second simple variant within complex variant (may repeat for 3rd simple variant, etc.)				
A	CWE	48008-7	Simple variant:		1.1	38485^NM_000106.5(CYP2D6):c.1457G>C (p.Ser486Thr)^ClinVar	
			<u>Transcript specification</u>				
C	CWE	48019-4	DNA sequence variation type		1.1	LA6690-7^Substitution^LN	SNV
D	CWE	48018-6	Gene:		1.1	2625^CYP2D6^HGNC	
E	CWE	51958-7	Transcript Reference Sequence ID (aka NM_RefSeq):		1.1	NM_000106.5^^RefSeq	
F	CWE	41103-3	DNA change:		1.1	c.1457G>C^^c.HGVS	
G	CWE	48005-3	Amino acid change:		1.1		

2016-05-09 -5PM OutlineGenomicModel_ for HL7 - Lite Genetics report

				<u>Genomic specification</u>			
H		48013-7		Genomic Reference Sequence:		NG_008376.3^^RefSeqGene	
I	ST	69547-8		Genomic Reference (Ref) allele:		G	
J	NM	81254-5		Genomic Allele location:		42126611	
K	ST	69551-0		Genomic Alternate (Alt) allele:		C	
				<u>Other optional codes related to simple variation</u>			
L	CWE	48004-6		dbSNP:		rs368949613^^dbSNP	
M	CWE	81256-0		COSMIC			
N	CWE	81257-8		CIGAR			
				<u>Other possible attributes</u>			
O	NM	81258-6		Allelic Frequency NFR		0.59168	
P	CWE	48001-2		Chromosome region		22q13.2	
				<u>Allelic state and interpretive attributes</u>			
Q	CNE	53034-5		Allelic state:		LA6706-1^Heterozygous^LN	
R	CNE	53037-8		Clinical significance:		LA6668-3^Pathogenic^LN	
S	TX	81259-4		Associated phenotype:		688395015^Debrisoquine adverse reaction (disorder)^SCT	

Report Part 4 – Structural (copy number) variations

	Data Type	LOINC #	R C O	Observation display Name- draft version	card	OBX-4	OBX-5 Example values	Comments
		81297-4		Structural (Copy number) variants section – panel- Note some structural variants can be reported via strict ClinGen IDs and/or COSMIC IDs in the above sections				
A	CWE	81286-7	P	Structural (copy number) variant I			nsv995237^17p12(chr17:14184616-15581544)x1	
C	NR	81287-5	P	Structural variant reported start-end:			14184616-15581544	
D	??	81288-3	O	Precision of boundaries:			??	Needs ideas regarding how to express

2016-05-09 -5PM OutlineGenomicModel_ for HL7 - Lite Genetics report

E	NM	81299-0	C	Structural variant reported arrCGH ratio:				Simple number (believe less than 1- will have to check) Only available if the method type is arrCGH
F	CWE	48019-4	P	DNA sequence variation type:			LA6686-5^Duplication^LN	
G	NM	81300-6	O	Structural variant length:			1,396,929	Don't often see this detail in clinical reports
H	NR	81301-4	O	Structural variant outer start-end:				Don't often see this detail in clinical reports
I	NR	81302-2	O	Structural variant inner start-end:			14184616-15581544	Don't often see this detail in clinical reports
J	CWE	81290-9	O	Structural variant HGVS:			NC_000017.11:g.(?_14184616)_(15581544 _)dup	
K	CWE	81291-7	C	Structural variant ISCN:				Include if available
L	CWE	81298-2	P	Structural variant cytogenetic location:			17p12	
M	CWE	81304-8	P	Structural variant method type			Sequencing	Answer list to be developed but would include FISH, arrCGH, MLFP, sequencing, next generation sequencing (and maybe CNV-seq, FREEC, readDepth, CNVnator, SegSeq and event-wise testing (EWT)). See PMID:23527109/PMCID: PMC3604020