**GSCID/BRC Clinical Metadata Standard**

**v1.2  
Finalized by the GSCID/BRC Clinical Metadata Working Group**

How to interpret the document:

BOLD: Field name   
ITALICS: Attributes of the field  
  
Clinical metadata may include elements that pose a risk of patient identification. These have been marked for each entry in an attribute labeled 'Risk'. Individuals handling these data should familiarize themselves with the relevant regulations regarding the sharing and use of protected health information. As specified in the NIAID contracts to the GSCs and BRCs, it is the institutions’ responsibility to provide the appropriate privacy training for handling this information. Fields marked as HIPAA data elements are based on the information at the HIPAA web site which describes what protected health information includes. For dates: "All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older." For geographic location: "All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes" are considered potentially identifiable.

Several required fields are dates (e.g., sample collection date) which should be supplied in the international standard (ISO) format (yyyy-mm-dd). However, these dates when combined may potentially lead to patient identification. To avoid this possible breach of patient health information, the dates should be collected but censored from public access. In place of dates, time relative to the sample collection date should be provided. In the case of multiple samples at different dates from the same individual, then the initial sample collection date should be used as the reference point. For example, a patient with a birth date of 2010-01-01 who was hospitalized 2010-10-01 and had blood drawn for follow-up study 2011-01-01 would be shown as a 1 year old patient who was hospitalized 2 months prior to blood collection for the study. When providing the sample collection date, it should be de-identified by randomization. We recommend a randomization strategy with some linkage to seasonality since some of the infectious diseases are seasonal. The randomization process should be confined to a defined, relatively short temporal window, for example plus or minus 15 days.

Some fields such as those pertaining to Physical Exams, Diagnosis, Lab Tests, and Treatments may have multiple related entries. Preferred approaches for capturing multiple entries are described in the ‘Multiple Sets’ attribute.

We understand that LOINC codes provide information that cover multiple required fields (e.g., Lab Test Method; Body Site). However, not all data contributors will start with LOINC codes and we want to make sure that we capture needed information in a manner that could be easily mapped to LOINC codes to be used as standard values.

**1. Enrollment Criteria**   
*Name:* Enrollment Criteria   
*Multiple Sets:* This would be captured at the project level and should be a single common value for all enrolled participants   
*Risk:* Some risk if criteria are extremely restrictive   
*Definition:* The criteria used to identify and recruit human subject participants into a clinical research study, including both the inclusion and exclusion criteria, which describes the common characteristics of the base human population participating in the research study.   
*NA allowed:* No   
*Data Type:* text   
*Accepted Values:* free text   
*Examples:* Adults between the ages of 18 and 80 without evidence of previous exposure to the influenza A H5N1 virus   
*External reference:*:   
*Notes:*

**2. Unique ID**   
*Name:* Unique ID/ corresponds to Specimen Source ID in Core Sample metadata collection   
*Multiple Sets:*   
*Risk:* Should be non-traceable in the dataset to any individual person (i.e., de-identified). Considered high risk if any of the HIPAA named identifiers are used. Medical Record Numbers should never be used.   
*Definition:* Unique identification (ID) number that is assigned to a subject (patient, study participant, etc.) and is used to track the subject.   
*NA allowed:* No   
*Data Type:* alphanumeric   
*Accepted Values:* any unique string   
*Examples:* AC160   
*External reference:*:   
*Notes:* Just needs to be unique within the given study.

**3. Sample Collection Date**  
*Name:* Sample Collection Date/corresponds to Specimen Collection Date in Core Sample metadata collection   
*Multiple Sets:*   
*Risk:* Specific dates in general are an identification risk if tied to HIPAA-defined protected health information dates, such as admission date or date of death.   
*Definition:* Date the sample was collected from the subject.   
*NA allowed:* No   
*Data Type:* string   
*Accepted Values:* ISO format; Year (Required), Month and Date (Optional); yyyy-mm-dd.   
*Examples:* 2009-07-31   
*External reference:*:   
*Notes:* Core metadata. If the collection of a single sample occurs over an extended time, simply indicate the start date of the collection.

**4. Age**   
*Name:* Age/ corresponds to Specimen Source Age - Value and Specimen Source Age - Unit in Core Sample metadata collection   
*Multiple Sets:*   
*Risk:* Date of birth is a HIPAA identifier, age is also a HIPAA identifier when over 89   
*Definition:* Biological age of the subject (i.e. age of the subject at the time of sampling) at the time of collection or study enrollment.   
*NA allowed:* Yes   
*Data Type:* numeric   
*Accepted Values:* Year (Required), Month and Day (Optional), decimal acceptable when the age is less than 1 year; for age over 89, should use “age 90 or older”.   
*Examples:* 25   
*External reference:*:   
*Notes:*

**5. Gender/Sex**   
*Name:* Gender/Sex/corresponds to Specimen Source Gender in Core Sample metadata collection   
*Multiple Sets:*   
*Risk:*   
*Definition:* Gender or sex of the subject.   
*NA allowed:* No   
*Data Type:* string   
*Accepted Values:* male, female, unknown, missing   
*Examples:* male   
*External reference:*:   
*Notes:*

**6. Geographic Location**  
*Name:* Geographic Location/ corresponds to Specimen Collection Location- Location in Core Sample metadata collection   
*Multiple Sets:*   
*Risk:* Geographic location is a HIPAA identifier when smaller than a state.   
*Definition:* Country or geopolitical location where the sample was collected from the subject.   
*NA allowed:* No   
*Data Type:* text   
*Accepted Values:* Country Format: NCBI Country List, Geographic Location Vocabulary GAZ   
*Examples:* Uganda   
*External reference:*:   
*Notes:*

**7. Race/Ethnicity**   
*Name:* Race/Ethnicity   
*Multiple Sets:*   
*Risk:*   
*Definition:* Biological race, group or cultural background with which the subject most identifies with.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Anthro-Ethnicity-Groups list with codes   
*Examples:* Mixed, of predominantly African Ancestry 1301900   
*External reference:*:   
*Notes:* Derived from - SJ Mack, A. Sanchez-Mazas, D Meyer, RM Single, Y Tsai, HA Erlich. Anthropology/Human Genetic Diversity Joint Report. Chapter2: Methods Used in the Generation and Preparation of Data for Analysis in the 13th International Histocompatibility Workshop, in Immunobiology of the Human MHC: Proceedings of the 13th International Histocompatibility Workshop and Conference, J. Hansen, Editor. 2007, IHWG Press: Seattle. p. 564-579.

**8. Study type**  
*Name:* Study type   
*Multiple Sets:*   
*Risk:*   
*Definition:* Term(s) that describes the type or design of the conducted study.   
*NA allowed:* No   
*Data Type:* text   
*Accepted Values:* cross-sectional survey, cohort (prospective or retrospective), health facility, case-control, other   
*Examples:* prospective cohort   
*External reference:*:   
*Notes:*

**9. Health Status**  
*Name:* Health Status/ corresponds to Specimen Source Health Status in Core Sample metadata collection   
*Multiple Sets:*   
*Risk:* When related to "Specification of mother and fetal health status, should also include abortion"   
*Definition:* Impact of disease on subject function as reported by the patient at the time of physical examination.   
*NA allowed:* Yes   
*Data Type:* string   
*Accepted Values:* General finding of observation of patient terms in SNOMED   
*Examples:* asymptomatic   
*External reference:*: http://purl.bioontology.org/ontology/SNOMEDCT/118222006   
*Notes:* With reference to infectious disease being studied

**10. Comorbidity**   
*Name:* Comorbidity   
*Multiple Sets:* Multiple Comorbidities may be recorded during sample collection: comma delimited   
*Risk:*   
*Definition:* Additional disorders or diseases co-occurring with the primary disease; or the effect of such additional disorders or diseases. Primary disease: the disease caused by the suspected pathogenic organism   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Disease Terms ICD9/10, Disease Ontology   
*Examples:* acquired immunodeficiency   
*External reference:*:   
*Notes:*

**11. Concomitant Medication**   
*Name:* Concomitant Medication   
*Multiple Sets:* Multiple Concomitant Medication may be recorded during sample collection: comma delimited   
*Risk:*   
*Definition:* Medications being administered to or taken by the patient that are unrelated to the suspected pathogenic organism.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* RxNorm   
*Examples:* Ciprofloxacin   
*External reference:*: http://www.nlm.nih.gov/research/umls/rxnorm/   
*Notes:*

**12. Measured Attribute**   
*Name:* Measured Attribute   
*Multiple Sets:* Multiple Measured Attributes may be collected in the same Physical Exam process. Each Measured Attribute may have a Measured Method and Measured Value   
*Risk:*   
*Definition:* Quantitative measurements of the subject during the physical exam.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Vital signs terms in LOINC   
*Examples:* diastolic blood pressure 10 hour mean   
*External reference:*: https://loinc.org/   
*Notes:* Additional values can be defined via project-specific data dictionary

**13. Measured Method**  
*Name:* Measured Method   
*Multiple Sets:* One Measured Method is associated with one Measured Attribute   
*Risk:*   
*Definition:* Method used to capture the quantitative measurements during the physical exam.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Taking patient vital sign terms in LOINC   
*Examples:* temperature of skin palpation   
*External reference:*: https://loinc.org/   
*Notes:*

**14. Measured Value Name Measured Value**   
*Multiple Sets:* One Measured Value is associated with one Measured Attribute   
*Risk:*   
*Definition:* Actual results of lab physical exam recorded as numeric value, units recorded as text   
*NA allowed:* Yes   
*Data Type:* numeric, text   
*Accepted Values:* Units: metric, Scale: nominal, ordinal, quantitative, ordinal quantitative. Standard International (SI) units encouraged.   
*Examples:* 37 degrees C for temperature; 3 on the Blantyre Coma Scale (0-5)   
*External reference:*:   
*Notes:* Could be abstracted to normal/abnormal value

**15. Clinical Observation**   
*Name:* Clinical Observation   
*Multiple Sets:* Multiple Clinical Observations may be associated with the same Physical Exam   
*Risk:*   
*Definition:* Clinical observations that are not quantitatively measured during the physical exam.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* On examination - specified examination finding terms in SNOMED   
*Examples:* blood in vomit on examination   
*External reference:*: http://purl.bioontology.org/ontology/SNOMEDCT/271880003   
*Notes:*

**16. Relative Physical Exam Date**   
*Name:* Relative Physical Exam Date   
*Multiple Sets:* One Physical Exam Date may be associated with multiple Measured Attributes   
*Risk:* Specific dates in general are an identification risk if tied to HIPAA identifiers such as admission date or date of death.   
*Definition:* Time the physical exam assessment of the subject was performed relative to the date of the sample collection event.   
*NA allowed:* Yes   
*Data Type:* numeric   
*Accepted Values:* numeric; may include year-month-day-hour-sec   
*Examples:* 5 years and 6 months and 1 day   
*External reference:*:   
*Notes:*

**17. Disease Course**  
*Name:* Disease Course   
*Multiple Sets:*   
*Risk:*   
*Definition:* The totality of all processes through which a given disease instance is realized.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Acute disease onset, Chronic disease course, Chronic disease course with acute onset, Transient disease course, Transient disease course with acute onset, Progressive disease course, Progressive disease course with acute onset, Fatal disease course, Fatal disease course with acute onset   
*Examples:* fatal disease course   
*External reference:*: http://purl.obolibrary.org/obo/OGMS\_0000063   
*Notes:* Acute disease onset - a disease onset in which signs and symptoms progress relatively rapidly in comparison with a canonical disease course. Chronic disease course - A disease course that (a) does not terminate in a return to normal homeostasis and (b) would, absent intervention, fall within an abnormal homeostatic range. Chronic disease course with acute onset – a chronic disease course with an acute disease onset. Transient disease course - A disease course that terminates in a return to normal homeostasis. Transient disease course with acute onset – a transient disease course with an acute disease onset. Progressive disease course - A disease course that (a) does not terminate in a return to normal homeostasis and (b) would, absent intervention, involve an increasing deviation from homeostasis. Progressive disease course with acute onset – a progressive disease course with an acute disease onset. Fatal disease course – A progressive disease course that terminates in the death of the organism bearing the disease. Fatal disease course with acute onset – a fatal disease course with an acute disease onset.

**18. Symptom**   
*Name:* Symptom   
*Multiple Sets:* Multiple symptoms may be associated with the same patient. Each symptom may have an Onset Age, Onset Speed, and Severity   
*Risk:*   
*Definition:* Perceived change in function, sensation, loss, disturbance or appearance reported by the subject indicative of a disease.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* symptom terms: Symptom Ontology   
*Examples:* fever   
*External reference:*: http://www.ontobee.org/browser/index.php?o=SYMP   
*Notes:*

**19. Symptom Onset**  
*Name:* Symptom Onset   
*Multiple Sets:* Symptom Onset is associated with one Symptom   
*Risk:*   
*Definition:* One Symptom Onset is associated with one Symptom   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Human Phenotype Ontology (HPO): onset adult onset, childhood onset, congenital onset, infantile onset, juvenile onset, neonatal onset   
*Examples:* childhood onset   
*External reference:*: http://bioportal.bioontology.org/ontologies/49621/? p=terms&conceptid=HP%3A0003674   
*Notes:*

**20. Symptom Onset Speed**  
*Name:* Symptom Onset Speed   
*Multiple Sets:* One Symptom Onset Speed is associated with one Symptom   
*Risk:*   
*Definition:* The speed of the first appearance of symptoms of an illness. The actual time scale for onset will vary for different diseases and should be defined according to domain experts.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Human Phenotype Ontology (HPO): Speed of onset acute, chronic, insidious onset, subacute   
*Examples:* acute   
*External reference:*: http://bioportal.bioontology.org/ontologies/49684/?p=terms&conceptid=HP%3A0011008   
*Notes:*

**21. Symptom Severity**   
*Name:* Symptom Severity   
*Multiple Sets:* One SymptomSeverity is associated with one Symptom   
*Risk:*   
*Definition:* Severity, intensity, or variable expressivity of symptoms.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Human Phenotype Ontology (HPO); PATO Phenotype terms of intensity: Mild, Moderate, Severe   
*Examples:* moderate   
*External reference:*: http://bioportal.bioontology.org/ontologies/49712/?p=terms&conceptid=PATO%3A0000049   
*Notes:*

**22. Diagnostic Result**  
*Name:* Diagnostic Result   
*Multiple Sets:* Multiple Diagnostic Results may result from the same diagnosis process. However, there should be only one Primary Diagnostic Result that is about the presence of disease caused by the infectious agent under study.   
*Risk:*   
*Definition:* Diagnosis is the outcome of the assessment of a disease or injury, its likely prognosis and treatment.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Disease Terms ICD9/10, Disease Ontology   
*Examples:* malaria   
*External reference:*: http://www.ontobee.org/browser/index.php?o=DOID   
*Notes:*

**23. Relative Diagnosis Date**  
*Name:* Relative Diagnosis Date   
*Multiple Sets:*   
*Risk:* Specific dates in general are an identification risk if tied to HIPAA identifiers such as admission date or date of death.   
*Definition:* Time the diagnosis of the subject was determined relative to the date of the sample collection event.   
*NA allowed:* Yes   
*Data Type:* numeric   
*Accepted Values:* Numeric; may include year-month-day-hour-sec   
*Examples:* 5 years and 6 months and 1 day   
*External reference:*:   
*Notes:*

**24. Lab Test Type Detecting Pathogen**  
*Name:* Lab Test Type Detecting Pathogen   
*Multiple Sets:* Multiple Lab Tests may occur. One Lab Test Type Pathogen is associated with one LabTest.   
*Risk:*   
*Definition:* Long common lab test name that corresponds to the observation/phenomenon being measured to describe presence of specific pathogen.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Long common name as defined by LOINC/HL7   
*Examples:* 50556-0 (Urinalysis dipstick panel - Urine by Automated test strip)   
*External reference:*: https://loinc.org/   
*Notes:* \*Ref. Clinical Metadata: lab tests and treatment example metadata components

**25. Lab Test Type Pathogen Attribute**  
*Name:* Lab Test Type Pathogen Attribute   
*Multiple Sets:* Multiple Lab Tests may occur. One Lab Test Pathogen Attribute associated with one Lab Test   
*Risk:*   
*Definition:* Long common lab test name that corresponds to the observation/phenomenon being measured to describe attributes and/or severity of infection of a specific pathogen identified in Lab Test Type Detecting Pathogen.   
*NA allowed:* Yes   
*Data Type:* text, numeric   
*Accepted Values:* Long common name as defined by LOINC/HL7   
*Examples:* 468-9 (LOINC code for Sulfamethoxazole sensitivity test by MIC)   
*External reference:*: https://loinc.org/   
*Notes:* pathogen attributes and/or severity include drug resistance, inflammatory response, pathogen load, etc. Multiple rows for multiple resistance tests

**26. Body Site**  
*Name:* Body Site   
*Multiple Sets:* One Body site is associated with one LabTest   
*Risk:*   
*Definition:* Location on body or bodily substance/fluid from which materials used in lab tests were obtained from which the sample used in lab test is collected   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Terminologia Anatomica: hemisphere (if applicable) and site/subsite name or bodily substance/fluid name   
*Examples:* urine   
*External reference:*:   
*Notes:*

**27. Lab Test Method**  
*Name:* Lab Test Method   
*Multiple Sets:* One Lab Test Method is associated with one LabTest   
*Risk:*   
*Definition:* Primary method used in lab test to obtain measurements and attributes pertaining to the pathogen   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* LOINC: Primary method used in lab test name (e.g. immunoassay, PCR...)   
*Examples:* Minimum Inhibitory Concentration   
*External reference:*: https://loinc.org/   
*Notes:* \*Ref. Clinical Metadata: lab tests and treatment example metadata components. \*Row is redundant if using LOINC codes for LabTestTypePathogen and/or LabTestTypePathogenAttribute => optional

**28. Lab Test Measurement**  
*Name:* Lab Test Measurement   
*Multiple Sets:* One Lab Test Measurement is associated with one Lab Test.   
*Risk:*   
*Definition:* Actual results of lab test recorded as numeric value, units recorded as text   
*NA allowed:* Yes   
*Data Type:* numeric, text   
*Accepted Values:* Units: metric, Scale: nominal, ordinal, quantitative, ordinal quantitative. Standard International (SI) units encouraged.   
*Examples:* >0.5µg   
*External reference:*: https://loinc.org/   
*Notes:* \*Ref. Clinical Metadata: lab tests and treatment example metadata components

**29. Lab Test Standard**  
*Name:* Lab Test Standard   
*Multiple Sets:* One Lab Test Standard is associated with one LabTest   
*Risk:*   
*Definition:* Accrediting entity that the testing laboratory references/adheres to for procedural/measurement standards for microbial tests.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* CLSI; EUCAST; IFCC; ILAC; CLIA; ISO 15189 other - specify   
*Examples:* CLSI   
*External reference:*: http://www.clsi.org/ http://www.eucast.org/ http://www.ifcc.org/ http://www.iso.org/iso/catalogue\_detail?csnumber=56115 http://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/index.html?redirect=/clia/ https://www.ilac.org/   
*Notes:* EUCAST only applies for antibiotic resistance breakpoint standards \*\*\*\*\*When submitting metadata for antibiotic resistance users MUST include what standards were used to define S/R classification breakpoints if they are NOT submitting MIC numbers (CLSI and EUCAST S/R breakpoints are not comparable => resistance metadata in S/R format can only be meaningfully aggregated within breakpoint standards but not between them)

**30. Relative Lab Test Date**  
*Name:* Relative Lab Test Date   
*Multiple Sets:*   
*Risk:* Specific dates in general are an identification risk if tied to HIPAA identifiers such as admission date or date of death.   
*Definition:* Time the lab test of the subject was determined relative to the date of the sample collection event.   
*NA allowed:* Yes   
*Data Type:* numeric   
*Accepted Values:* numeric; may include year-month-day-hour-sec   
*Examples:* 5 years and 6 months and 1 day   
*External reference:*:   
*Notes:*

**31. Treatment Type Pathogen**  
*Name:* Treatment Type Pathogen   
*Multiple Sets:* Multiple Treatment Type Pathogen is associated with one/multiple Diagnosis and/or LabTest   
*Risk:*   
*Definition:* Treatment given to the subject directly related to the diagnosed pathogen   
*NA allowed:* Yes   
*Data Type:* drug/treatment   
*Name:* text, units: quantitative, body site: text, devices used: text   
*Accepted Values:* RxNorm: Drug dose (generic chemical name + method of delivery + units), Procedure (e.g excision + devices used),   
*Examples:* 309114 (cephalexin 500 MG Oral Capsule)   
*External reference:*: http://www.nlm.nih.gov/research/umls/rxnorm/   
*Notes:*

**32. Treatment Body Site**  
*Name:* Treatment Body Site   
*Multiple Sets:* One Body site is associated with one Treatment Type   
*Risk:*   
*Definition:* Location on body or avenue through which treatment was administered   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Terminologia Anatomica: hemisphere (if applicable) and site/subsite name   
*Examples:* oral   
*External reference:*: https://loinc.org/   
*Notes:* \*\*\*Row is redundant if using RxNorm codes for Treatment Type Pathogen when treatment only involves administration of drugs => optional. Row is NOT redundant if treatment involves surgical intervention => mandatory

**33. Relative Treatment Onset Date**  
*Name:* Relative Treatment Onset Date   
*Multiple Sets:*   
*Risk:* Specific dates in general are an identification risk if tied to HIPAA identifiers such as admission date or date of death.   
*Definition:* Relative time from the sample collection from the date when treatment of the subject started.   
*NA allowed:* Yes   
*Data Type:* text, numeric   
*Accepted Values:* numeric; may include year-month-day-hour-sec   
*Examples:* 5 years and 6 months and 1 day   
*External reference:*:   
*Notes:*

**34. Treatment Duration**  
*Name:* Treatment Duration   
*Multiple Sets:* Multiple Treatment Duration is associated with corresponding Treatment Type   
*Risk:*   
*Definition:* Time interval in which treatment took place. Possible ways of capturing data without violating HIPAA   
*NA allowed:* Yes   
*Data Type:* text, numeric   
*Accepted Values:* Time stamp (single point in time), Time interval (prolonged treatment and/or multiple instances of single treatment/dosage)   
*Examples:* 3/24hours   
*External reference:*:   
*Notes:*

**35. Treatment Outcome**  
*Name:* Treatment Outcome   
*Multiple Sets:* Multiple Treatment Outcome corresponding to completed rounds of Treatment Type   
*Risk:*   
*Definition:* Indication of whether treatment of the subject was effective at the end of the treatment duration. A successful treatment outcome is contingent on the patient not being re-diagnosed with an ailment associated with the targeted pathogen at the end of a specific Treatment Duration   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* successful / unsuccessful   
*Examples:* unsuccessful   
*External reference:*:   
*Notes:*

**36. Vaccine Type**  
*Name:* Vaccine Type   
*Multiple Sets:*   
*Risk:*   
*Definition:* The type of vaccine administered to the subject during the study.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Vaccine Ontology: Live attenuated, inactivated, subunit, toxoid, conjugated, DNA, recombinant vector based, recombinant protein   
*Examples:* live attenuated   
*External reference:*: http://www.violinet.org/vaccineontology/   
*Notes:*

**37. Vaccine Source**  
*Name:* Vaccine Source   
*Multiple Sets:*   
*Risk:*   
*Definition:* Name of the vaccine manufacture and vaccine production batch number or the lot number.   
*NA allowed:* Yes   
*Data Type:* text, numeric   
*Accepted Values:* Vaccine Ontology MERK, PENTACEL, NOVARTIS, GSK, MEDIMMUNE, etc. and batch/lot number as numeric and text   
*Examples:* GSK Flulaval, Batch No; 22011-13, Lot No:OH90019AB   
*External reference:*: http://www.violinet.org/vaccineontology/   
*Notes:*

**38. Vaccine Dosage**  
*Name:* Vaccine Dosage   
*Multiple Sets:*   
*Risk:*   
*Definition:* Number of times the subject was vaccinated during the study.   
*NA allowed:* Yes   
*Data Type:* numeric, text   
*Accepted Values:* integer   
*Examples:* 3   
*External reference:*: http://www.violinet.org/vaccineontology/   
*Notes:*

**39. Vaccine Dosage Amount**  
*Name:* Vaccine Dosage Amount   
*Multiple Sets:*   
*Risk:*   
*Definition:* The total amount (concentration or total volume) administered per each vaccination to subject during the study.   
*NA allowed:* Yes   
*Data Type:* numeric   
*Accepted Values:* value + SI unit pairs   
*Examples:* 0.5ml   
*External reference:*: http://www.violinet.org/vaccineontology/   
*Notes:*

**40. Vaccine Adjuvant**  
*Name:* Vaccine Adjuvant   
*Multiple Sets:*   
*Risk:*   
*Definition:* Component that potentiates the immune responses to a vaccine antigen and enhance the desired immune responses.   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Vaccine Ontology: Vaccine adjuvant Nonadjuvanted, adjuvanted with alum, MF59, monophosphoryl lipid A (MPL) etc.   
*Examples:* nonadjuvanted   
*External reference:*: http://purl.obolibrary.org/obo/VO\_0000580   
*Notes:*

**41. Relative Vaccination Date**   
*Name:* Relative Vaccination Date   
*Multiple Sets:*   
*Risk:* Specific dates in general are an identification risk if tied to HIPAA identifiers such as admission date or date of death.   
*Definition:* Time of vaccination(s) of the subject relative to the date of the sample collection event.   
*NA allowed:* Yes   
*Data Type:* numeric   
*Accepted Values:* numeric; may include year-month-day-hour-sec   
*Examples:* 5 years and 6 months and 1 day   
*External reference:*: http://www.violinet.org/vaccineontology/   
*Notes:* \*need to think about versioning, backwards compatibility

**42. Site of Vaccination**   
*Name:* Site of Vaccination   
*Multiple Sets:*   
*Risk:*   
*Definition:* The mode in which vaccine was administered to the subject. SC = Subcutaneous, IM = Intramuscular, IV = Intravenous, IN = Intranasal, O = Oral   
*NA allowed:* Yes   
*Data Type:* text   
*Accepted Values:* Vaccine Ontology: Route of administration SC, IM, IV, IN, Oral   
*Examples:* im   
*External reference:*: http://www.ontobee.org/browser/rdf.php?o=VO&iri=http://purl.obolibrary.org/obo/VO\_0000574   
*Notes:*