Minimum Data Fields for Non-human Mammal Surveillance

Finalized January 21, 2009 by DMID/NIAID
Updated April 30, 2010
Updated March 24, 2011
Updated November 12, 2015

This document contains the required and optional minimum data fields for non-human mammal surveillance. The NIAID CEIRS centers must send surveillance data to the Influenza Research Database (www.fludb.org) in the format detailed below. Fields will be evaluated and revised on a semi-annual basis.

How to interpret the document:

**BOLD:** Name of field

**ITALICS:** Definition of field, format for values of this field, and curation to be carried out on data.

**REQUIRED FIELDS**

1. **Collector Name (REQUIRED)**
   Definition: The name of the person collecting the sample. If needed, separate multiple names by comma.
   Format: Text (“Unknown” if name is not known)
   Curation: None

2. **Collecting Institution (REQUIRED)**
   Definition: Code of the institution to which the person collecting the sample belongs. Codes will be assigned by NIAID/IRD.
   Format: Text
   Curation: None

3. **Collection Date (REQUIRED if available)**
   Definition: Date on which the field sample was collected. Some samples will be sent to the centers and not collected by the centers. Therefore, sometimes the laboratory will not have a collection date. **If the collection date and received date are both available, then both should be reported in two separate fields.** Either field #3 or #4 must be provided. Provide both #3 and #4 if available.
   Format: (d)d-first 3 letters of month-yyyy or unknown
   Curation: Leading 0 of dd need not be provided; years must have four digits; month must match the first three (3) letters of the month; checked for out-of-range dates. This ensures that there is no ambiguity with formats of dates from different regions of the world.
4. Receipt Date (REQUIRED if available)
   
   **Definition:** Date on which the field sample was received in the laboratory. Some samples will be sent to the centers and not collected by the centers. Therefore, there needs to be the option to choose either collection date or receipt date. **If both dates are available, then both this field and #3 should be reported in two separate fields.** Either field #3 or #4 must be provided.

   Provide both #3 and #4 if available.

   **Format:** (d)d-first 3 letters of month-yyyy or unknown

   **Curation:** Leading 0 of dd need not be provided; years must have four digits; month must match the first three (3) letters of the month; checked for out-of-range dates. This ensures that there is no ambiguity with formats of dates from different regions of the world.

5. Location (REQUIRED)
   
   **Definition:** Descriptor of local sampling environment, such as name of wild life refuge, name of zoo, name of retail market. Include the name of the city. Examples: St. Marks National Wildlife Refuge, St. Marks

   **Format:** Text

   **Curation:** None

6. Country (REQUIRED)
   
   **Definition:** Country in which original sample was collected

   **Format:** Value must be found on the approved list used by NCBI and found at the NCBI web site: http://www.ncbi.nlm.nih.gov/projects/collab/country.html

   **Curation:** Must appear in NCBI country codes as above

7. State/Province (REQUIRED)
   
   **Definition:** Descriptor of sampling location that is one gradation finer than country, for example Minnesota, Alberta, Guangxi. Must spell out the name of the state/province

   **Format:** Text

   **Curation:** Must appear in list of approved names; combined list from all CEIRS maintained and updated by IRD

8. Host Species (REQUIRED)
   
   **Definition:** Full **scientific name** of host genus and species, without abbreviations or unknown. If genus is known but species is unknown, then use the genus followed by "sp." If both the genus and species are unknown, then use the scientific family name, without abbreviation. The Integrated Taxonomic Information System (ITIS) should be used for identifying the scientific name (http://www.itis.gov/).

   **Format:**
   - Text
   - Use U = Unknown (without family name)
   - Use Env(ironment) for samples taken from the environment (e.g., a water sample or feces picked up from the beach) when the source species is not known.

   **Curation:** Currated against ITIS. The scientific name must be provided (not the common name)

   **Example:** Sus scrofa
9. **Host Common Name (REQUIRED)**  
   *Definition:* The English common name given to a particular species, without abbreviations or unknown.  
   *Format:* Text  
   *Curation:* Curated against ITIS.  
   *Example:* Wild boar

10. **Host Identifier (REQUIRED)**  
    *Definition:* An identifier assigned to the host by the collector. If multiple samples are collected from a host, then all must be linked to the same host identifier. If band number is part or all of the host ID, identify the banding agency in optional field #3, Host ID Type.  
    *Format:* Free form. Up to 50 characters of any combination of alphanumeric characters and special characters except: single quote, double quote, ampersand, hash (pound sign), forward slash, carriage returns, question mark, equal or plus signs, greater than or less than signs, percent and period. Leading and trailing spaces are also forbidden.  
    *Curation:* None

11. **Capture Status (REQUIRED)**  
    *Definition:* Reason for sample collection  
    *Format:* A=Active surveillance (e.g. trap); sampling not initiated by an outbreak; live capture  
               K=Opportunistically sampled (e.g., hunter-killed, salvage or killed in order to take sample)  
               O=Sample collected in response to outbreak and/or disease investigation  
               M= Market sample  
               P= Sample obtained from a pet  
               Z= Sample obtained from a zoo  
               F= Farm sample  
               AB=Abattoir  
               SEH=Semi-enclosed housing  
               OT= Other; plus free text (e.g. OT- restaurant)  
               U=Undetermined/Unknown  
    *Curation:* Must be one of the allowed values or text. Text entries will not be curated.

12. **Health (REQUIRED)**  
    *Definition:* If Capture Status is A, O, M, P, Z, F, OT or U, then health status of animal at time of sample collection. If Capture Status is K or AB, then Health is U(undetermined/Unknown)  
    *Format:* H(healthy)/S(ick)/D(ead)/U(ndetermined)  
    *Curation:* Must be one of the allowed values.

13. **Age (REQUIRED)**  
    *Definition:* Age of animal from which sample taken  
    *Format:* W/J/A/P/S/U (abbreviations for Weanling, Juvenile, Adult, Pup, Subadult or Undetermined/Unknown)  
    *Curation:* Must be one of the allowed values.

14. **Gender (REQUIRED)**
15. Behavior (REQUIRED)
   Definition: Classify the sampled animal’s natural habit as one of the following:
   • Wild: indisputably wild, such as sea lions sampled at a rehab center.
   • Captive-Wild: captive wild animals such as tigers at zoos
   • Domestic: indisputably domestic, such as pigs on pig farms
   • Domestic feral: Domestic animal which is feral, such as feral pigs.
   • Unknown
   • Other: those that cannot be readily placed into above categories.
   Format: W(ild)/CW(captive-wild)/D(omestic)/DF(domestic-feral)/U(known)/text. Text will be used to describe the “Other” category.
   Curation: Must be one of the allowed values or text. Text entries will not be curated.

16. Sample Identifier (REQUIRED)
   Definition: Identifier assigned to each sample collected. If multiple samples are taken from the same host, each sample should have its own identifier.
   Format: Free form. Up to 50 characters of any combination of alphanumeric characters and special characters except: single quote, double quote, ampersand, hash (pound sign), forward slash, carriage returns, question mark, equal or plus signs, greater than or less than signs, percent and period. Leading and trailing spaces are also forbidden.
   Curation: None

17. Sample Material (REQUIRED)
   Definition: Material on which the testing was performed. If multiple samples are taken from the same host, they must be entered as separate records but if taken from one host they should be linked to that host (see item 8 above).
   Format:
   • OP = Oral-pharyngeal swab
   • NS = Nasal swab
   • TS = Tracheal swabs
   • RS = Rectal swab
   • B = Blood
   • SR = Serum
   • TT = Tissue from trachea
   • TB = Tissue from brain
   • TO = Other tissue
   • F = Feces
   • L= Lungs
   • BAL = Bronchio-alveolar lavage
   • LL = Lung lavage fluid
   • S = Soil
• W = Water
• SL = Slurry
• O = Other; plus free typed out text (e.g. O – spleen)
• U = Unknown

_Curation:_ Must be one of the allowed values or text. Text entries will not be curated. Can provide more detailed information through the other category. Can provide more detailed information through the other category (e.g. O – spleen).

18. Test for influenza virus (REQUIRED)

_Definition:_ Method by which the CURRENT assessment of a sample as flu-positive/negative is made. If multiple tests were performed and the site prefers to report them all now, repeat fields 18-19. For PCR-based tests, the relevant target genomic segment should also be described. Further testing can be reported with the “viral isolate data” fields.

_Format:_ Text

_Curation:_ None

_Example:_ RT-PCR H5, antigen ELISA, AGID or other flu test

19. Influenza test result (REQUIRED)

_Definition:_ Classification of sample as flu positive or negative based on the test performed and reported in field #18.

_Format:_ P/N

_Curation:_ Must be one of the allowed values.

**REPEAT Fields #18-19 if multiple tests were performed and the site prefers to report them all now. Further testing can be reported with the “viral isolate data” fields.

20. Other pathogens tested for (other than influenza) (REQUIRED)

_Definition:_ Classification of samples tested for organisms other than influenza virus such us PRRS virus or PCV virus, etc.

_Format:_ Text

_Curation:_ None

21. Other pathogen test result (other than influenza) (REQUIRED)

_Definition:_ Classification of sample as positive or negative or unknown based on the test performed and reported in field #20.

_Format:_ P/N/U

_Curation:_ Must be one of the allowed values.

**REPEAT Fields #20-21 if multiple tests for other pathogens were performed.**

**OPTIONAL FIELDS**
1. **Latitude (OPTIONAL)**  
   *Definition:* Latitude by GPS of site at which original sample was collected. If not available, a course estimate is acceptable.  
   *Format:* WGS84, decimal degrees, signed, five digit accuracy when available  
   *Curation:* Checked for within-range

2. **Longitude (OPTIONAL)**  
   *Definition:* Longitude by GPS of site at which original sample was collected. If not available, a course estimate is acceptable.  
   *Format:* WGS84, decimal degrees, signed, five digit accuracy when available  
   *Curation:* Checked for within-range

3. **Host ID Type (OPTIONAL)**  
   *Definition:* The type or format and agency assigning the host Identifier (Required field #12).  
   *Format:* Text  
   *Curation:* None  
   *Example:* Patuxent BBL leg band, Moscow banding center leg band, web tag, UAF hunter id, UAF unique name, etc. (see [http://www.pwrc.usgs.gov/BBL/homepage/aboutaux.cfm](http://www.pwrc.usgs.gov/BBL/homepage/aboutaux.cfm) for additional types of markers and [http://www.pwrc.usgs.gov/BBL/homepage/link.cfm](http://www.pwrc.usgs.gov/BBL/homepage/link.cfm) for information on international banding schemes)

4. **Transport_Medium (OPTIONAL)**  
   *Definition:* Medium within which sample is transported from field to lab  
   *Format:*  
   - VTM = Viral transport medium  
   - G (guanidine)  
   - E (ethanol)  
   - Other – use text to describe  
   *Curation:* Must be one of the allowed values or text. Text will not be curated.

5. **Strain Name (OPTIONAL)**  
   *Definition:* It is recommended that a Strain_Name ONLY be provided when there is a viral isolate, otherwise use “Unknown”. In other words, “Unknown” should be utilized for the Strain_Name field when a sample is PCR+ but there is no viral isolate. The subtypes determined by PCR, sequencing, etc. are to still be reported in the dedicated HA_subtype and NA_subtype fields (i.e. fields 9-12 in viral isolate data). The format, A/chicken/Fujian/411/2002(HxNx), will be used in cases where a partial subtype has been determined by viral isolation (*For example, A/chicken/Fujian/411/2002(HxN1)). In cases where there are mixed subtypes contained within a sample use A/chicken/Fujian/411/2002(mixed) for Strain_Name although, these multiple subtypes can be reported in dedicated HA_subtype and NA_subtype fields (i.e. fields 9-12 in viral isolate data).  
   *Format:* Text  
   *Curation:* Must be WHO strain name, Unknown, A/chicken/Fujian/411/2002(HxNx), or A/chicken/Fujian/411/2002(mixed)
6. **Other Host and Sample Data (OPTIONAL)**

   - **Definition:** Text describing anything else of interest.
   - **Format:** Text
   - **Curation:** None

*Examples: A/Viet Nam/1203/2004 (H5N1)*