Diagnostic tests

STRUCTURE OF THE DATA AVAILABLE

Field name	Data type	Cleaning notes
refID	Numerical	Unique identification of a reference.
FullReference	Free-text	Full reference in the format: "all authors, YEAR, publication title,journal,issue, pages".
groupID	Numerical	This field identify UNIQUE STUDY GROUPS within each paper.
context	Categorical	The type of study being conducted, 4 options were possible in the RADIO list: cohort; cross sectional; diagnostic test validation; and longitudinal.
year	Numerical	The year the study was conducted, when reported. -1 when not reported.
agent	Categorical	VBD agent. Coded by EFSA. Collected in the form as a RADIO LIST.
agentSubtypeDC1	Categorical	Only used in the first phase of data collection (DACRAH1), listed some expected subtypes: AHS1, BTV2, BTV23, BTV4, BTV8, BTV9.
targetSpecies	Categorical	Collected in the form as a RADIO LIST including all species eligible for the review.
ageDC1	Categorial	In the first project of data collection (DACRAH1), this was collected as categories – Young or Adult. To avoid problems with the definition, or groups with ranges of ages, in DACRAH2 this was substituted by the field below.
ageMonths	Numerical	Age of the animals in the study, when reported. -1 when not reported.
infMode	Categorical	"Natural infection" or "Experimental infection".
sampStrategy	Categorical	Sampling strategy. Collected in the form as a RADIO LIST.
sampUnit	Categorical	Sampling Unit. Collected as a radio list, and during data cleaning the adta entry was checked and corrected when not used consistently (for instance when "single" was used to mean "animal"). In the end, only "animal" or "herd/flock" were listed.
sampUnitSize	Numerical	Sample unit size.
route	Categorical	Route of infection, when reported. (For cases where animals where purposefully infected to test diagnostic sensitivity).
dose	Free-text	Description of dose, in case of animals purposefully infected.
intervention	Categorical	"Index test" or "reference test" (Status of evaluated test). The gold standard test is the *reference test* i.e. the one that they know has positive results; the novel test is the *index test* or one which is being validated.
targetLab	Categorical	Target of the diagnostic tests: antibody, antigen or nucleic acid.
labtest	Categorical	Set list of diagnostic tests, based on the tests approved for use in the European Union for each of the VBDs considered.
labTest_C	Free-text	Any additional comments about the laboratory test above.
reagents	Free-text	Any additional comments about the reagents used in the laboratory test above.
testLimit	Free-text	Any additional comments about the cut-off limits used in the laboratory test above.
timePoint	Numerical	Days since the animals were purposively infected, when that was the case, in longitudinal studies.
nTested	Numerical	Number of animals tested
nTruePositive	Numerical	Number of true positive animals, when known.
nTrueNegative	Numerical	Number of true negative animals, when known.
nPositive	Numerical	Number of positive animals using the test being evaluated.
nNegative	Numerical	Number of negative animals using the test being evaluated.

sensitivity	Numerical (percentage)	Reported as percentage (for instance 90 for 90%, NOT 0.9). Rarely given, so left BLANK when not given.
specificity	Numerical (percentage)	Reported as percentage (for instance 90 for 90%, NOT 0.9). Rarely given, so left BLANK when not given.
UCI_Sen	Numerical (percentage)	Upper Control Interval of the Confidence Interval for the sensitivity.
LCI_Sen	Numerical (percentage)	Lower Control Interval of the Confidence Interval for the sensitivity.
UCI_Spe	Numerical (percentage)	Upper Control Interval of the Confidence Interval for the specificity.
LCI_Spe	Numerical (percentage)	Lower Control Interval of the Confidence Interval for the specificity.
crossReactivity	Free-text	Any cross reactivity detected should be reported.
rowID	Numerical	A unique number for all rows in the dataset.
uniqueID	Numerical	A unique number for all study groups in the entire dataset (combination of refID + groupID).
matrix	Categorical	The matrix samples for diagnostic. Blood or Serum only were reported in the papers reviewed, but many other categories are available.
studyType	Categorical	Created during data cleaning, a variable to help identify if the study compared the index test to a reference test ("comparison" category), reported only an index text - "index", or only a reference test – "reference").
ShortBibliography	Free-text	Reference in the format "First author, et al. YEAR".
Author	Free-text	List of authors
Title	Free-text	Publication title
Abstract	Free-text	Abstract
publicationYear	Free-text	Publication year.

NOTES AND WARNINGS ON DATA MEANING AND INTERPRETATION, ASSUMPTIONS AND SHORTCOMINGS

General notes about the data collection format:

1) Data rows with the same refID are results reported from the same study
2) Individual study groups within these references receive the same groupID. The groupID can therefore be used to track results that refer to the same animal group.
3) Combinations of refiD+ groupID represent UNIQUE animal groups for which results are reported. These two fields should be used to identify multiple rows of outcomes that refer to the same animal group (during data cleaning we combined these two and created "uniqueID").
4) Data collection is performed in Distiller using "data collection forms". Each form results in one row when the data are looked in the tabular format (for instance in Excel of .CSV format). Every output can only be reported once in each form, therefore to report multiple values of the same type of outcome for the same group (say tests performed in different time points), the entire form must be duplicated. *This is really important to consider when generating data summaries for specific parameters – the possibility of duplicated information within each study group needs to be carefully considered and accounted for (remove duplicates per group).*

Notes about this research objective in particular:

1) Please note that the number of positive and negative animals should be considered in light of the true number of positive and negative animals, when reported. Another possibility, is that the results of a "reference test" and an "index test" were reported for the SAME animal group (groupID). In this case, the results of the index test should be interpreted *in comparison* to the reference test.
2) Sensitivity and specificity were only documented when *reported in the paper*. The data collectors have not calculated sensitivity and specificity from the results given in a paper themselves.