D3.3: Recommendations on the usefulness of current methodologies for providing standardised outputs from disease control programmes and identified knowledge gaps

E. Meletis[1], B. Conrady[2, 3], P. Hopp[4], T. Lurier[5, 6], J. Frössling[7, 8], T. Rosendal[7], C. Faverjon[9], L. P. Carmo[10], J. J. Hodnik[11], L. Ózsvári[12], P. Kostoulas[1], G. van Schaik[13, 14], M. Nielen[13], T. Knific[15], J. Schulz[16], S. Šerić-Haračić[17], A. Madouasse[18]

[1] Department of Public and One Health, School of Health Sciences, University of Thessaly, Karditsa, 43100, Greece

[2] Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Frederiksberg, Denmark

[3] Complexity Science Hub Vienna, Vienna, Austria

[4] Norwegian Veterinary Institute, POB 64, NO-1431 Ås. Norway

[5] Université de Lyon, INRAE, VetAgro Sup, UMR EPIA, F-69280 Marcy l'Etoile, France

[6] Université Clermont Auvergne, INRAE, VetAgro Sup, UMR EPIA, F-63122 Saint-Genès-Champanelle, France

[7] Department of Disease Control and Epidemiology, National Veterinary Institute (SVA) – Uppsala, Sweden

[8] Department of Animal Environment and Health, Swedish University of Agricultural Sciences – Skara, Sweden

[9] Ausvet Europe, Lyon, France

[10] Veterinary Public Health Institute, Vetsuisse Faculty, University of Bern, Bern, Switzerland

[11] Clinic for Reproduction and Large Animals-Section for Ruminants, Veterinary faculty, University of Ljubljana, Ljubljana, Slovenia

[12] Department of Veterinary Forensics and Economics, University of Veterinary Medicine Budapest - Budapest, Hungary

[13] Department of Population Health Sciences, Unit Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University – Utrecht, the Netherlands

[14] Royal GD – Deventer, the Netherlands

[15] Veterinary Faculty, Institute of Food Safety, Feed and Environment, University of Ljubljana, Ljubljana, Slovenia

[16] Institute of Epidemiology, Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Greifswald - Insel Riems, Germany

[17] University of Sarajevo Veterinary faculty Sarajevo, Bosnia and Herzegovina

[18] BIOEPAR, INRAE, Oniris – Nantes, France

Funded by the Horizon 2020 Framework Programme of the European Union

One of the main conclusions of the work undertaken in working group 3 is that the context of the prevalence of infection matters when choosing a method for quantifying the probability of freedom from infection. When infection is absent from an area, the objective is to prove that it has not been introduced to secure trade with partners outside the area. The level of interest is in this case the whole area. The scenario tree methodology and more recent simulation methods are well suited to this context (Martin et al., 2007a; b; Rosendal et al., 2020). When the infection is still endemic, as in the initial phase of an eradication programme, the objective is to identify herds that are free from the infection within the programme to secure trade within and outside the area. In this case, historical data from the surveillance programme can be used to make inference and enhance the identification of infected herds. The STOC free model was designed to operate in this context (Madouasse et al., 2022). Below, we provide more detail on these ideas.

The assumption made in most work on proving disease freedom is that the disease is truly absent from the region or country of interest. The methodological questions revolve around quantifying the evidence of absence from the absence of evidence. Because it is impossible to quantify uncertainty from a hypothesis of complete absence, while this uncertainty exists, most methods quantify the probability of obtaining a negative surveillance outcome if the disease were present at a chosen (low) prevalence called the *design prevalence*.

A historical perspective on the methods developed for quantifying the probability of freedom from infection is of interest to understand the state-of-the-art. Assuming disease absence from a region or a country, early work therefore focused on determining a sample size to prove that the infection prevalence was not greater than a chosen design prevalence with a

certain level of confidence. This was initially done assuming homogeneous populations, in which all the animals had a similar probability of being infected, and a single perfect test was performed for detection of infection. These assumptions were later relaxed by considering the imperfect sensitivity of the tests and the fact that animals are usually clustered within farms (Hanson et al., 2003; Verdugo et al., 2015;). The impact of imperfect specificity is usually neglected, because any positive test result will be investigated (diagnostic follow-up) until either proven to be a false positive or confirmed positive. In this sense, the proportion of false positives is reduced to 0, yielding a perfect *Sp*. A later refinement was the inclusion of differences in the probability of infection between different animals or herds allowing for the estimation of probabilities of freedom from infection from surveillance systems that relied on risk-based sampling (FAO, 2014). Risk-based sampling permits to increase the effectiveness of surveillance by focusing surveillance efforts on areas where the infection is more likely to be found. The **scenario tree method** was designed to estimate a probability of freedom from infection and surveillance sensitivity from complex surveillance data with differences in the probability of infection in different components of the surveillance system (risk-based sampling) as well as imperfect sensitivity of the testing procedure (imperfect test sensitivity, sampling, hierarchical structure of the data with animals nested within farms). This, together with the fact that simulations could be run on spreadsheets led to this method being widely used for substantiating freedom from infection (Norström et al., 2014). More recent simulation models use the same principles as the scenario tree methodology.

When the infection is absent from the area under investigation, it has to be (re-introduced to be present. Typically, this can be considered as infection emergence and should be a rare event with people making efforts to prevent the introduction of the infection through

the routes of emergence perceived as important. Incorporating a probability of introduction into a model of infection freedom is therefore a difficult task. Probabilities are usually thought of as the frequency of a certain event of interest under a large number of instances where that event could happen, such as the frequency of heads in a large number of coin flips. It is almost impossible to estimate a probability of introduction from data, as there are no or not enough similar cases of disease introduction in similar contexts. In this case, probabilities can be conceived as beliefs, which may have a rational basis, about the probability of introduction. The first view represents a frequentist perspective on probabilities and the second a Bayesian perspective. This is important to reflect upon when designing models and communicating their outcome. More generally, when the infection is absent, the whole estimation process relies on simulating its presence or its introduction under different scenarios, and out of these, count the proportion of times this infection would be detected by the surveillance system. Therefore, all the methods considered can be seen as conceptualising the problem of substantiating freedom from infection as a Bayesian problem in which what is evaluated is the probability of the hypotheses, notably in the form of a design prevalence, given the surveillance data collected. This explains why in some papers, the probability of the infection being present before collecting the data is referred to as the prior and the estimated probability of infection freedom as the posterior.

When the infection is still endemic, data from infected herds can be used to estimate strengths of association between risk factors and the probability of infection, thereby improving the detection of infection in herds not yet detected. This is the approach followed in the Bayesian models proposed by Heisey et al., (2014) and Madouasse et al., (2022). By exploiting the correlation in longitudinal test results, the model by Madouasse et al., (2022) also estimates

the herd-level sensitivity and specificity of the tests used in the surveillance programme as well as the monthly probabilities of getting and eliminating the infection. By making inference from surveillance data, these models are less reliant on hypotheses whose validity can be hard to assess and provide predictions that are adapted to the context in which surveillance is performed. Such models can also produce knowledge that is transferable to other surveillance systems. However, when the infection is absent or rare, there is no added value to these Bayesian inference and prediction models since in those cases, they will perform simulations from the prior distributions used as input.

The definition of what is an infected herd can also be difficult to formulate. For example, for bovine viral diarrhoea virus, there are different types of infected animals that do not pose the same epidemiological risk and have different diagnostic test results. Persistently infected (PI) animals are the main source of infection for other animals. They shed massive amounts of virus but do not produce antibodies. Transiently infected animals shed the virus for a few weeks and produce antibodies against the virus for many years. Unborn PI foetuses are epidemiologically important, but it is not currently possible to detect infection before birth. The definition of what makes an infected herd in a surveillance programme could include or exclude either of the latter two categories. Translating sensitivity or specificity at the animal level into their equivalent at the herd level can therefore be a real challenge. The issue can be difficult to deal with when using latent class models. With these models, the latent class, that should correspond to the definition of what the epidemiologists mean by infected herds, will depend on the prior distributions used for the sensitivity and the specificity of the different diagnostic tests used. As these sensitivities and specificities can be unknown and modelled using best guesses, the model may work with a latent class that is different from the intended definition

of an infected herd. The same problem also exists with simulation-based models that use sensitivities elicited from experts. However, misspecification of model parameters is impossible to detect because, in the worst case, the infection will emerge once and be missed which can always be considered compatible with a surveillance sensitivity that is lower than one. The definition of sensitivities and specificities at the herd level is therefore a research gap that needs to be filled. A modelling framework needs to be developed to estimate these parameters from animal data and the required data need to be identified.

The objective of any method that quantifies a probability of freedom from infection from surveillance data is (i) to evaluate the confidence in a country's freedom from infection claim or (ii) to assist stakeholders in making decisions about which countries or herds to trade with. The output of such methods should therefore be understandable and usable by stakeholders. Following (Cameron, 2012), the most reported outputs are the sensitivity of surveillance as well as the probability of freedom from infection, which can be easily obtained using the scenario tree method. The STOC free model returns a posterior distribution for the probability of infection which can be translated into a distribution for the probability of freedom from infection. However, translating statistical distributions into information that is usable by decision makers can be challenging. Further work on communicating the output of quantitative models to decision makers should address this question. In addition to the specific output from methods for assessment of disease freedom, there is also a need to put this result into context and to make the information that will support decision more complete by adding qualitative aspects. Guidelines for incorporating necessary information about surveillance attributes have been proposed e.g. by RISKSUR (Peyre et al., 2019) and AHSURED (Comin et al., 2019).

Outputs related to the costs of error, especially the cost of declaring an infected country as free from infection, have been proposed. Although the cost of different surveillance programme designs can be estimated (Meyer et al., 2019), the cost of the consequences of an undetected emerging infectious disease are harder to predict. A first step in this direction consists in estimating the time to detection and the size of the outbreak at detection using models incorporating population dynamics and disease spread as proposed by (Rosendal et al., 2020).

Another concern is the fact that when using any method, an assumption is that the available data accurately reflect what they are supposed to measure. This assumes an infrastructure that collects these data in a reliable way and that the modellers are aware of all the limitations of the data they have, and the outcomes of the model can relatively easily be interpreted to the 'users' (e.g. farm managers, traders, veterinary officials).

Lastly, our work focused on cattle diseases. Most of the concepts and methods reviewed could be applied to other animal and plant species to quantify probabilities of absence of specific diseases in these species or even absence of these species. On the other hand, although few studies in fields not related to cattle farming were included in this work (e.g., Ramsey et al. 2009), future methodological development may benefit from considering more work conducted in other fields. There is a large number of scientific publications on the problem of evaluating the probability of absence of a species after an eradication programme. Sometimes, these species are problematic because they were introduced in a new environment in which they caused significant damage (Ramsey et al., 2009; Rout et al., 2014). At other times, the focus is on estimating the probability that a human disease has been eradicated following a

vaccination campaign. Although the problems may be different, the focus of these studies is on quantifying the evidence of absence from the absence of evidence and could therefore provide interesting ideas for such problems in cattle, and more broadly in livestock.