

A rationale for assessing the effectiveness of animal disease surveillance systems

Timothée Vergne^{1*}, Barbara Häslér^{1,2}, Julian Drewe¹, Vladimir Grosbois³

¹ Veterinary Epidemiology, Economics and Public Health group, Royal Veterinary College, London, United Kingdom

² Leverhulme Center for Integrative Research on Agriculture and Health, Royal Veterinary College, London, United Kingdom

³ Animal et Gestion Intégrée des Risques group (Research Unit 22), Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Montpellier, France

*tvergne@rvc.ac.uk

Keywords: Surveillance, evaluation, effectiveness, type I and II errors, decision-making

Abstract

Surveillance systems produce data which, once analysed and interpreted, support decisions regarding disease management. While several performance measures for surveillance are in use, no theoretical framework has been proposed yet with a rationale for defining and estimating effectiveness measures of surveillance systems in a generic way. An effective surveillance system is a system whose data collection, analysis and interpretation processes lead to decisions that are appropriate given the real disease status of the target population. Accordingly, we developed a framework accounting for sampling, testing and data interpretation processes, to depict in a probabilistic way the direction and magnitude of the discrepancy between “decisions that would be made if the true state of a population was known” and the “decisions that are actually made upon the analysis and interpretation of surveillance data”. The proposed framework provides a theoretical basis for standardised quantitative evaluation of the effectiveness of surveillance systems. We illustrate such approaches using a hypothetical surveillance system aimed at monitoring the prevalence of an endemic disease. Finally we discuss the potential of this new approach for harmonising cost-effectiveness analyses.

Introduction

Surveillance has been defined as “the systematic measurement, collection, collation, analysis, interpretation and timely dissemination of animal health and welfare data from defined populations whose data are essential for describing health hazard occurrence and to contribute to the planning, implementation and evaluation of risk mitigation actions” [1]. As a consequence, surveillance systems can be viewed as observation systems whose function is to provide evidence to inform decisions regarding disease management [2,3]. Therefore, evaluations of the effectiveness of a surveillance system should be done in relation to the decisions made upon the analysis and interpretation of the data generated by surveillance [3]: does the surveillance system generate data that lead to a decision that would have been made if the true state of the population was known?

The data produced by a surveillance system are generated through reporting, sampling, sample testing processes and diagnosing. Reporting and sampling are rarely exhaustive and may be non-representative. Moreover, diagnostic and sample testing procedures may misclassify a fraction of the tested samples. Therefore, data generated by surveillance systems are most of the time imperfect. The characterization of the epidemiological status of the

population produced through the analysis and interpretation of such data nonetheless informs decisions.

To date, the evaluation of the effectiveness of surveillance systems was done by estimating a wide range of measures such as the number of days from introduction to detection, the number of infected units detected by the surveillance, the sensitivity of the detection, etc. The use of these various variables limits standardisation and comparability between surveillance programmes. In this paper, we propose a rational probabilistic approach that overcomes these limitations.

The objectives of the present paper are 1) to present a conceptual framework defining a generic measure of effectiveness, 2) to illustrate how this measure can be used with the example of a hypothetical surveillance system aimed at monitoring prevalence of an endemic disease and 3) to discuss how this measure can be extended to other surveillance objectives and used by decision-makers when designing surveillance systems.

Conceptual framework to estimate the effectiveness of a surveillance system

We define the effectiveness of a surveillance system as its ability to generate an observation of the population that is in accordance with the true disease status of the population so that the decisions that are made based on this observation are those that would have been made if the true status of the population was known. Therefore, in order to draw a conceptual framework for the assessment of the effectiveness of surveillance systems, we need to clearly define a *population state variable* that is relevant depending on the objectives of the surveillance and the *observations* generated by the surveillance system. Once these are defined, effectiveness measures can be estimated.

The state variable

The relevant state variable is defined according to the objectives of the surveillance. A surveillance system aiming at monitoring prevalence may have a different state variable than a surveillance system aiming at early detection. The state variable can be categorised into different subsets based on the decision rules related to interventions. For example, the simplest categorisation would be:

- S+: the subset of possible states that would require to make a decision about a possible intervention,
- S-: the subset of possible states that would not require to make a decision

For a surveillance system aiming at monitoring prevalence, the state variable is the true prevalence of the disease of

interest. S+ can be defined as the subset of states where the true prevalence is above a pre-defined threshold (the status is non-acceptable), and S- as the subset of states where the true prevalence is below this threshold (the status is acceptable). State subsets can also be defined for other surveillance objectives (e.g., early detection, demonstrating freedom from a disease or case finding).

The observation

A surveillance system produces data that is analysed and interpreted in order to evaluate the state of the population. Here, “observation” refers to the evaluation of the state of the population based on the data produced by the system. The “observation” results from the generation of surveillance data and the analysis and interpretation of these data to inform a decision based on the same thresholds used to define the subsets of states. As for the state variable, possible observations can be categorised into different subsets based on the decision rules. In the simple example introduced above, observations can be categorised into the two following subsets:

- O+: the subset of possible observations that lead to the making of a decision about a possible intervention,
- O-: the subset of possible observations that do not lead to the making of a decision.

There should be an equal number of subsets of observations and of subsets of states. For the surveillance system aiming at monitoring prevalence, O+ could be defined as the subset of observations where the apparent prevalence is above a pre-defined threshold (the same threshold that has been used to define the subsets of states) and O- as the subset of observations where the apparent prevalence is below this threshold. This example is clearly over-simplistic, but we believe this is necessary to understand the logic of this rationale. Further complexity will be discussed later.

The effectiveness measures

When S+, S-, O+ and O- are specified, two types of errors can be defined in the line of statistical or diagnostic tests (Table 1).

Table 1: Classification of the different types of errors for a surveillance system

| | S+ | S- |
|----|---------------|--------------|
| O+ | | Type I error |
| O- | Type II error | |

In our example, the risk of type I error is thus $\Pr(O+|S-)$, i.e. the probability that the surveillance system produces information that lead to the making of a decision when the real state of the population would not require it. Type I errors imply that decisions are unnecessarily made leading to unnecessary expenditures.

Similarly, the risk of type II error is defined as $\Pr(O-|S+)$, i.e. the probability that the surveillance system produces information that does not lead to the making of a decision although the true state of the population would require it. Type II errors imply that no decisions are made despite a genuine health issue occurring.

We propose to evaluate the effectiveness of a given surveillance system by estimating the probability of a type I error and the probability of a type II error.

Estimating the effectiveness measures

To estimate the probabilities of type I and type II errors, three different steps are required:

- 1) To describe the thresholds determining the subsets of true population states associated with different decisions to be made,
- 2) To describe the systematic data collection processes that give rise to the surveillance data, including sampling, reporting, diagnostic and testing,
- 3) To describe the data analysis and interpretation processes that determine the subsets of observations associated with the different decisions to be made.

Given these three steps have been completed, the probabilities of type I and type II errors can be estimated either analytically (for simple case studies) or using simulations (in more complex situations).

Application to a hypothetical surveillance system aiming at monitoring prevalence

This conceptual approach is illustrated with a hypothetical surveillance system aiming at monitoring the prevalence of a disease to inform decision-makers who will then use this information to decide whether or not to implement an intervention (for this example, we assumed that no other source of information is used to make this decision). We followed the three steps defined in the previous paragraph. Step 1: The designers of the system and the decision-makers are assumed to have defined two subsets of population states:

- S+ is the subset of the states of the population where the true prevalence p is above or equal to 0.2. The decision to implement interventions is required.
- S- is the subset of the states of the population where the true prevalence p is below 0.2. The decision to implement interventions is not required.

Step 2: The surveillance system to be evaluated is described as precisely as possible: it consists of a random sampling of $n=100$ population units thorough the year (the population is considered infinite) and testing of these units using a perfect diagnostic test.

Step 3: Investigating the data analysis and interpretation processes that lead to decisions regarding the implementation of interventions: the decision to implement interventions is made whenever the proportion of positive samples (n_p/n) the previous year is above or equal to 0.2. Therefore:

- O+ is the subset of observations where the proportion of positive tests the previous year is above or equal to 0.2. The decision to implement interventions is made.
- O- is the subset of observations where the proportion of positive tests the previous year is below 0.2. The decision to implement interventions is not made.

Evaluating the effectiveness of this surveillance system requires estimating the probability of making an appropriate or an inappropriate decision regarding the implementation of interventions. The quantities of interest are therefore:

- $\Pr(O|S^-) = \Pr(n_p < 0.2n | p < 0.2)$
- $\Pr(O+|S+) = \Pr(n_p \geq 0.2n | p \geq 0.2)$
- $\Pr(O-|S+) = \Pr(n_p < 0.2n | p \geq 0.2)$
- $\Pr(O+|S-) = \Pr(n_p \geq 0.2n | p < 0.2)$

For this hypothetical situation, computing these quantities can be done analytically because the probability distribution of n_p is known: it is a binomial distribution of parameter n (the sample size) and p (the real prevalence of the disease in the population). So $\Pr(X \leq n_p < Y | p, n)$ can be computed for any value of X , Y , p and n , and this probability can be plotted against p (Figure 1).

In Figure 1, the black curve represents, for any value of p (true prevalence), the probability that, in view of the data produced by the surveillance system, the decision-makers decide to not implement interventions. The red curve represents its complement, *i.e.* the probability, for any value of p , to decide to implement interventions.

As shown in Figure 1, for a prevalence slightly below 0.2 (situation requiring no interventions), the hypothetical surveillance system generates data that, given the decision rule used, leads to an inappropriate decision (implementing interventions) with a probability of around 0.4 (probability of type I error). For a prevalence around 0.5, the interventions will be implemented with a probability of almost 100% (appropriate decision), and therefore the probability of type II error is almost null.

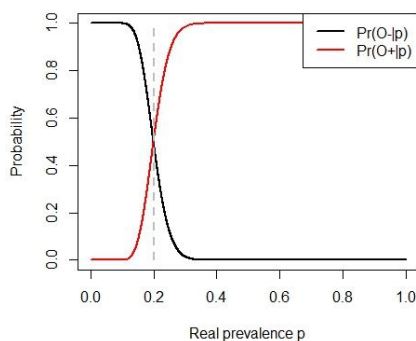


Figure 1: Probability of not implementing interventions (black curve) and of implementing interventions (red curve) as a function of the prevalence of the disease for our hypothetical surveillance system. The dashed grey line represents the threshold defined by decision makers to distinguish the states of the focal population not requiring (on the left hand side of the line) and requiring (on the right hand side of the line) the implementation of interventions.

Discussion

This paper presents a sound theoretical framework to estimate the effectiveness of surveillance systems.

This conceptual framework was illustrated with an unsophisticated surveillance system to monitor prevalence and a simple decision rule for the implementation of intervention if the prevalence increased above a certain threshold. It has to be highlighted that this approach can

easily be extended to more complex surveillance systems (several surveillance components, non-random sampling, imperfect diagnostic tests, etc.) and to more complex decision rules (where data interpretation accounts for imperfect test, where the decision is based on the precision of the estimation, where more than two subsets of states and observation are considered, etc.) Resolving these problems analytically soon becomes intractable. Using simulations provides a convenient and flexible alternative. Note that making the decision to implement control measures is likely to involve other types of considerations than only the surveillance data. Therefore the decision rule used in the example is likely to be over-simplistic but it was chosen for the sake of clarity. As soon as the decision process can be formalised (whatever its complexity), it can be incorporated within this framework.

Our application estimated the probabilities of type I and type II errors for a given surveillance system. We argue that this approach can be extended to compare alternative surveillance systems, with the objective to identify the most effective one. For example, the same approach has been used to assess the effectiveness of a surveillance system similar to the one described above with the exception that it samples 1000 units instead of 100. For a real prevalence slightly below 0.2, it was shown that such a surveillance system is associated with a much lower probability of type I error (it collapses to almost zero).

In this paper, we developed the framework for a surveillance system aiming at monitoring prevalence. The same approach can be generalized to surveillance systems with other surveillance objectives (early detection, freedom from disease and case finding). For example, this concept can be used to estimate the probability that a given surveillance system aiming at demonstrating freedom of a given disease will lead to the decision that the territory is free of the disease although it should not (type I error), and the probability that the surveillance system cannot demonstrate freedom although the territory is actually free (type II error).

The concept proposed provides a foundation for a standardised use of effectiveness in cost-effectiveness analysis. The next step is to assess the economic consequences for each type of error by estimating disease costs and surveillance and intervention costs. Therefore, for a given context and surveillance objective, the effectiveness of the system (as defined in this paper) and its economic value can be estimated, and sensitivity analyses can be used to establish which changes in the surveillance system allow increasing or decreasing the probabilities of type I and type II errors.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7 / 2007-2013) under grant agreement n° 310806.

References

1. Hoinville LJ, et al. (2013) *Prev Vet Med* 112: 1-12.
2. Häsler B, et al. (2011) *BMC Health Services Research* 11: 225.
3. Howe KS, et al. (2013) *Epid and Inf* 141: 91-101.