

# Development of a user-friendly and analytically strong platform to support near real-time surveillance of food animal diseases

ARRUDA AG<sup>1</sup>\*, MUELLNER P<sup>2</sup>, MUELLNER U<sup>2</sup>, AHLSTROM C<sup>2</sup>, MORRISON R<sup>1</sup>, PEREZ A<sup>1</sup>  
<sup>1</sup>University of Minnesota, St Paul, United States of America; <sup>2</sup>Epi-interactive, Wellington, New Zealand

\*arrud002@umn.edu

## Abstract

The objective of this project was to develop and demonstrate a user-friendly, analytically strong platform to provide insights into food animal disease occurrence, spread and control using porcine reproductive and respiratory syndrome (PRRS) and the Swine Health Monitoring Project (SHMP; United States national volunteer project) as a model.

An application composed of four dashboards was developed using the RStudio Shiny platform; fictitious data were used for demonstration purposes. The dashboards correspond to (i) visualisation of PRRS incidence over time; (ii) molecular characterisation of PRRS virus (PRRSV) isolates; (iii) PRRS space-time analysis; and (iv) network analysis of pig movement.

This innovative surveillance platform integrates disparate sources of data and provides insights into important aspects of PRRS epidemiology and, potentially, other food animal infectious disease spread and occurrence over time. A few examples of deliverables include monitoring of disease trends and near real-time identification of potential virus emergencies, construction of phylogenetic trees for comparison of PRRS virus isolates across swine farms, construction of risk maps to identify high-risk areas across the country, and ability to trace-back and identify important swine farms for surveillance purposes using pig movement. Importantly, it allows end-users, such as veterinarians, managers, and researchers, to interact with the database and explore it in a flexible manner to ultimately improve the quality of their decisions.

**Keywords:** *Epidemiological analytical tool, food animal diseases, real-time surveillance*

## Introduction

Swine diseases have shown an epidemic pattern globally over the years (1), with a few examples being the emergence of highly pathogenic PRRS viruses (2; 3), porcine epidemic diarrhea virus (PED; 4); and the Senecavirus A (5). There are multiple complex factors involved in those emergence and re-emergence processes; to name a few of them, the evolution of the pathogen itself, possible multiple transmission routes, current dynamics of the swine industry, and absence of regulated control strategies.

Control programs for PRRS are relatively common in North America, and serve for the purpose of data collection (6) for a

wide variety of objectives, including description of molecular diversity of viruses and creation of summaries that can aid in disease control and prevention. One example of such programs in the United States is the SHMP, which is a national voluntary project that currently enrolls approximately 42% of the sow population in the United States, with close to 1,000 swine sites. This program collects weekly data on occurrence of selected swine diseases, such as PRRS, PED, and Senecavirus A.

Those monitoring programs commonly capture and update information such as farm demographics, location, and diagnostic data on a regular basis. Additionally, further information such as animal movement and virus sequencing data is at times available from other sources (e.g. farm management software and diagnostic laboratories). Combining those disparate sources of surveillance data, as well as summarising and interpreting this large amount of information in real-time and on a meaningful and useful manner, is challenging.

That growing increase in the amount and quality of such available information can be paired with analytical advancements of epidemiological methods (e.g. network analysis and infectious disease modelling). That combination results in an opportunity to develop an integrated approach to visualise, interpret, and use surveillance data to inform decisions and improve the response to potential emerging diseases that could have a high economic impact for food animal industries. It is required that those tools are able to manage complex and large data, often referred to as Big Data, and to translate results into a user-friendly interface that fits the purpose of informing end-users. Primary users of such a platform are veterinarians, farm managers, and swine producers in charge of making decisions on disease management.

The objective of this project was to develop and demonstrate a surveillance data visualisation platform (PIG/savi [Shiny Analysis and Visualisation Insights for Swine Health management]) that is able to capture the different levels of complexity of the swine animal industry and provide useful insights into disease occurrence and potential spread and control using PRRS and the SHMP as a model.

## Materials and methods

PIG/savi was developed using the RStudio Shiny platform (available at <http://shiny.rstudio.com>). PIG/savi allows for online assessment as well as an offline capability by using

RStudio. Codes for each dashboard were wrapped into an R function that was called through the server-side script of the R Studio Shiny application. JavaScript and CSS were used to extend the functionality and visual elements of core RStudio Shiny. PIG/savi, in its current state, is composed of four dashboards that are demonstrated using fictitious data. Even though fictitious data is being used, it resembles the data that is received on a weekly basis through the SHMP, which is currently kept and updated in the format of Excel® spreadsheets (2010; Microsoft Corp., Santa Rosa, CA).

A description of the four PIG/savi dashboards (Figure 1) is provided as follows:

1. Disease frequency: allows for visualisation of disease trends over time for participating swine sites; consists of a cumulative incidence and an aggregate prevalence graph and allows for the ability to filter by attributes of interest (e.g. production system). This dashboard was created using the JavaScript library nvD3 (7).
2. Molecular characterisation of PRRSV isolates: using a segment of the PRRS genome (ORF5 gene), this dashboard allows for the visualisation of a homology matrix commonly used by field veterinarians (sequences aligned using the MUSCLE algorithm) and the construction of two different types of phylogenetic trees: neighbor-joining (NJ) and maximum likelihood (ML); as well as the ability to color tree nodes according to attributes of interest (e.g. year of isolation and production system). R packages used include ‘adegenet’ (8), ‘phangorn’ (9), and ‘ape’ (10).
3. Space-time analysis: this dashboard was created using the JavaScript library ‘Leaflet’ and the R package ‘splanets’ (11) and allows for visualisation of swine sites in a satellite map with changing colors representing different PRRS status over time, as well as the construction of a kernel smoothed incidence risk map.
4. Network analysis: the last dashboard allows for the visualisation of animal movements in and out of each participating swine site (Figure 2), as well as a data table option that show results from the calculation and extraction of farm-level network statistics. The R package used for this output was EpiContactTrace (12).

## Results and discussion

The tool developed herein is analytically strong and offers insights that cannot be provided by visualisation-only available tools. It centralises several important parts of the process of interpretation and investigation of spread of PRRS and potentially other infectious diseases, including spatial and temporal occurrence of outbreaks, genetic assessment of the pathogen’s genome, and implications of animal movements in disease transmission.

First, the disease frequency dashboard allows for monitoring of disease trends over time and the identification of disease seasonality trends as well as of potential epidemics, which can serve as an early warning surveillance system. This could be described across systems and regions, which could aid in

identification of different management practices as well as regional characteristics (e.g. environmental factors) that have an effect in the occurrence of PRRS outbreaks. This dashboard can further function as a surveillance tool for the evaluation of the efficacy of specific industry strategies implemented over time that target disease control and/or elimination.

The second dashboard presents a user-friendly view of methods that are commonly used at a local level to discriminate between novel and preexisting PRRS virus strains in order to infer spread patterns (6), which are the homology matrix and the NJ tree. This phylogenetic tree has traditionally been provided by North American diagnostic laboratories to field veterinarians and it informs decisions even though this distance matrix method produces only one tree and the reader cannot be confident of its robustness (13). For this reason, the tool also offers the option to construct a ML tree, which has the advantage of considering the probability that the given tree gave rise to the observed data (13). The introduction of this new method, in combination with what has traditionally been used, provides the end-user with a unique opportunity to take advantage of the different molecular methods to inform decisions without relying solely on one of them.

The spatial-temporal dashboard with the construction of risk maps allows for the identification of high-risk areas over time, which can help characterise risk and provide insights on how PRRS is spreading across the country. In addition, it may inform decisions such as movement of negative/ positive pigs to certain areas.

Finally, the animal movement dashboard allows for rapid trace back. This may be particularly useful not only for outbreak investigations when one wishes to rule out potential infection due to animal movement, but also when it is needed to presume status for other farms for the purposes, for example, of informing those who need to know during a control and eradication project. The advanced analytics provided by this dashboard could inform future risk-based surveillance strategies by aiding in the identification of key swine sites to sample (e.g. the ones with the highest out-degree values, or so-called “high-consequence” farms) in cases where funding needs to be prioritised.

In conclusion, the platform developed in this project aims to centralise, integrate, and analyse surveillance information from multiple sources to provide immediate and eventually near real-time insights into specific diseases. PIG/savi is being used on multiple diseases involving multiple animal species, and it can also potentially provide data for larger research projects in a standardised format. The tool allows users to interact with the database and explore it in a flexible manner to improve their decision making process.

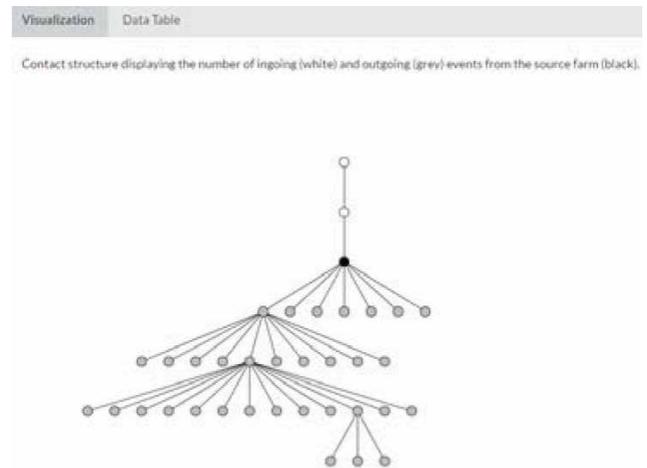
Organisation of such data and its interpretation in near real-time are pre-requisites for improving the speed of diagnosis of emerging infectious diseases, as well as their further

spread and prevention. For PRRS, the platform can support disease investigations, where molecular characterisation and trace-back play an important role. For those reasons, PIG/savi fits into the concept of near real-time surveillance, which is essential for timely responses to emerging diseases as exemplified above.

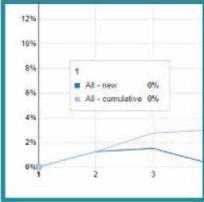
References

1. Meng XJ. *Transb. Emerg. Dis.* 59(1), 85-102, 2012
2. Workman *et al.* *Genome Announc.*, 4(4), e00772-16, 2016
3. Zhou *et al.* *J. Virol.* 86(11), 6373-6374, 2012
4. Stevenson *et al.* *J. Vet. Diagn. Invest.* 25(5), 649-654, 2013
5. Leme *et al.* *Transb. Emerg. Dis.* 62, 603-611, 2015
6. Perez *et al.* *JAVMA* 246(12), 1304-1317, 2015
7. nvD3. <http://nvd3.org/>
8. Jombart T. *Bioinformatics* 24, 1403-1405, 2008
9. Schliep KP. *Bioinformatics* 27(4), 592-593, 2011
10. Paradis *et al.* *Bioinformatics* 20: 289-290, 2004
11. Rowlingson B, Diggle PJ. *Geogr. Syst.* 2, 307-317, 1993
12. Nöremark M, Widgren S. *BMC Vet. Res.* 10, 71, 2014
13. Vandamme AM. *The Phylogenetic Handbook*, pp23-27, 2009

**Figure 2.** Selected details of one of PIG/savi’s dashboards – the animal movement and trace-back dashboard, showing farm of interest (black), incoming (white) and outgoing (grey) movements.



**Figure 1.** Screenshot showing the four dashboards currently available in PIG/savi.



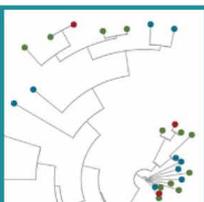
**Frequency**  
Explore the weekly and cumulative incidence and see aggregate prevalence of sow herd status.

EXPLORE



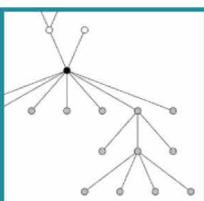
**Space Time Analysis**  
Navigate a map of PRRS herd status over time.

EXPLORE



**Genetics**  
Explore the genetic relationship of PRRS in different herds.

EXPLORE



**Movement**  
Explore contact networks between herds.

EXPLORE