PLOS ONE

Can biosecurity and network properties predict disease diversity in the salmonid industry?

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research Article
Full Title:	Can biosecurity and network properties predict disease diversity in the salmonid industry?
Short Title:	Disease diversity in salmonid farms in Ireland: the interplay between biosecurity and farm centrality
Corresponding Author:	Tadaishi Yatabe University of California Davis Davis, California UNITED STATES
Keywords:	Bayesian modeling; network analysis; generalized Poisson regression; risk-based surveillance; aquaculture; Ireland
Abstract:	Salmonid farming in Ireland is mostly organic, which implies limited disease treatment options, and its network of live fish movements possesses characteristics that would facilitate infection spread processes. This highlights the importance of biosecurity for preventing the introduction and spread of infectious agents. In this paper we characterized the biosecurity of the salmonid farms in Ireland using a survey, then developed a score for benchmarking the disease risk of salmonid farms. The usefulness and validity of this score - together with farm indegree, for predicting disease diversity, defined as the number of different diseases affecting a farm during a year, was assessed through generalized Poisson regression using Bayesian inference. Seawater salmon (SW salmon) farms had the highest biosecurity scores with a median (IQR) of 82.3 (5.4), followed by freshwater salmon (FW salmon), 75.2 (8.2), and freshwater trout (FW trout) farms 74.8 (4.5). For FW salmon and trout farms, the top ranked model (in terms of leave-one-out information criteria, looic) was the null model (looic = 46.1). For SW salmon farms, the best ranking model was the full model (looic = 33.3). Farms with a higher biosecurity score were associated with lower disease diversity, and farms with a high number of suppliers (indegree > 1) were associated with increased disease diversity. The effect of the interaction between these variables was also important, showing an antagonistic effect, which would indicate that biosecurity effectiveness is achieved through a broader perspective on the subject, which includes a minimization in the number of suppliers and hence in the possibilities for infection to enter a farm. The work presented here could be used to elaborate indicators of a farm's disease risk based on its biosecurity score and number of suppliers, to better inform risk-based disease surveillance and control strategies for private and official stakeholders.
Order of Authors:	Tadaishi Yatabe
	Simon J More
	Fiona Geoghegan
	Catherine McManus
	Ashley E Hill
	Beatriz Martinez-Lopez
Opposed Reviewers:	
Additional Information:	
Question	**Response**
Financial Disclosure	This study was funded by the Consejo Nacional de Inovacion, Ciencia y Tecnologia (Becas Chile scholarship to TY), http://www.conicyt.cl/; the University of California Davis (School of Vet Med GSSP award to TY), http://www.vetmed.ucdavis.edu/research/grad_education/gssp/; and Marine Harvest

Please describe all sources of funding that have supported your work. **This information is required for submission and will be published with your article, should it be accepted.** A complete funding statement should do the following: Include **grant numbers and the URLs** of any funder's website. Use the full name, not acronyms, of funding institutions, and use initials to identify authors who received the funding. **Describe the role** of any sponsors or funders in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. If the funders had **no role** in any of the above, include this sentence at the end of your statement: "*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*" However, if the study was **unfunded**, please provide a statement that clearly indicates this, for example: "*The author(s) received no specific funding for this work.*" * typeset	Ireland (TY), http://marineharvestireland.com/. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Competing Interests You are responsible for recognizing and disclosing on behalf of all authors any competing interest that could be perceived to bias their work, acknowledging all financial support and any other relevant financial or non-financial competing interests. Do any authors of this manuscript have competing interests (as described in the PLOS Policy on Declaration and Evaluation of Competing Interests)? **If yes**, please provide details about any and all competing interests in the box below. Your response should begin with this statement: *I have read the journal's policy and the authors of this manuscript have the following competing interests:* **If no** authors have any competing interests to declare, please enter this statement in the box: "*The authors have*	The authors have declared that no competing interests exist.

declared that no competing interests exist." * typeset	
Ethics Statement You must provide an ethics statement if your study involved human participants, specimens or tissue samples, or vertebrate animals, embryos or tissues. All information entered here should **also be included in the Methods section** of your manuscript. Please write "N/A" if your study does not require an ethics statement. **Human Subject Research (involved human participants and/or tissue)** All research involving human participants must have been approved by the authors' Institutional Review Board (IRB) or an equivalent committee, and all clinical investigation must have been conducted according to the principles expressed in the Declaration of Helsinki. Informed consent, written or oral, should also have been obtained from the participants. If no consent was given, the reason must be explained (e.g. the data were analyzed anonymously) and reported. The form of consent (written/oral), or reason for lack of consent, should be indicated in the Methods section of your manuscript. Please enter the name of the IRB or Ethics Committee that approved this study in the space below. Include the approval number and/or a statement indicating approval of this research. **Animal Research (involved vertebrate animals, embryos or tissues)** All animal work must have been conducted according to relevant national and international guidelines. If your study involved non-human primates, you must provide details regarding animal welfare and steps taken to ameliorate suffering; this is in accordance with the recommendations of the Weatherall report, "The use of non-human primates in research." The relevant guidelines followed and the committee that approved	N/A

the study should be identified in the ethics statement. If anesthesia, euthanasia or any kind of animal sacrifice is part of the study, please include briefly in your statement which substances and/or methods were applied. Please enter the name of your Institutional Animal Care and Use Committee (IACUC) or other relevant ethics board, and indicate whether they approved this research or granted a formal waiver of ethical approval. Also include an approval number if one was obtained. **Field Permit** Please indicate the name of the institution or the relevant body that granted permission.	
Data Availability PLOS journals require authors to make all data underlying the findings described in their manuscript fully available, without restriction and from the time of publication, with only rare exceptions to address legal and ethical concerns (see the PLOS Data Policy and FAQ for further details). When submitting a manuscript, authors must provide a Data Availability Statement that describes where the data underlying their manuscript can be found. Your answers to the following constitute your statement about data availability and will be included with the article in the event of publication. **Please note that simply stating 'data available on request from the author' is not acceptable.** ***If, however, your data are only available upon request from the author(s), you must answer "No" to the first question below, and explain your exceptional situation in the text box provided.*** Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?	Yes - all data are fully available without restriction
Please describe where your data may be found, writing in full sentences. **Your answers should be entered into the box below and will be published in the form you provide them, if your manuscript is**	All data set files will be available in a Dryad repository, upon article's acceptance.

accepted. If you are copying our sample text below, please ensure you replace any instances of XXX with the appropriate details. If your data are all contained within the paper and/or Supporting Information files, please state this in your answer below. For example, "All relevant data are within the paper and its Supporting Information files." If your data are held or will be held in a public repository, include URLs, accession numbers or DOIs. For example, "All XXX files are available from the XXX database (accession number(s) XXX, XXX)." If this information will only be available after acceptance, please indicate this by ticking the box below. If neither of these applies but you are able to provide details of access elsewhere, with or without limitations, please do so in the box below. For example: "Data are available from the XXX Institutional Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data." "Data are from the XXX study whose authors may be contacted at XXX." * typeset	
Additional data availability information:	Tick here if the URLs/accession numbers/DOIs will be available only after acceptance of the manuscript for publication so that we can ensure their inclusion before publication.

Davis, April 20th 2015

To whom it may concern,

The Irish salmon farming is mostly (80% + of the production) organic, and hence farmers have a very limited therapeutic arsenal at their disposal to treat disease. Additionally, in a recent publication we have showed that Ireland's network of live fish movement possesses characteristics that would facilitate infectious disease spread. These two features highlight the importance of biosecurity for preventing disease outbreaks, which could undermine the sustainability of this economic sector, vital to the rural western seaboard of Ireland.

In this research article we have characterized the biosecurity of the Irish salmonid farming industry using a survey approach, and based on it we have defined a biosecurity score for benchmarking farms. The score was validated through estimating its effect in fish health outcomes at the farm level, described as disease diversity, while accounting for the effect of farm network centrality (indegree, or the number of fish suppliers of a farm).

This is the first time the biosecurity of fish farms in Ireland has been characterized and, to the best of our knowledge, the first time the effect biosecurity (and a network centrality measure) on disease diversity has been estimated in a veterinary setting. For this, we have used Bayesian modeling through Hamiltonian Markov chain sampling, using Stan, an increasingly popular methodology, due to its ability to deal with complex posterior distributions. Model results have been extensively explored through counterfactual model predictions and simulation, displayed using appealing and easy to understand plots. In doing so, we have found the existence of a very meaningful interplay between biosecurity, farm network centrality, and disease.

We believe that these findings will useful for improving the biosecurity of the entire industry, allowing the benchmark of farms according to their disease risk, which in turn will help in the design and implementation of targeted disease surveillance and control plans within Ireland. Additionally, the methods described here could be applied to other animal production settings and countries, helping the decision making process of private and official stakeholders at a global scale.

Sincerely,

Tadaishi Yatabe
DVM, MPVM, PhD(c)

Can biosecurity and network properties predict disease diversity in the salmonid industry?

Tadaishi Yatabe*[1], Simon J. More[2], Fiona Geoghegan[3], Catherine McManus[4], Ashley E. Hill[5], Beatriz Martínez-López[1]

[1] Center for Animal Disease Modeling and Surveillance (CADMS), Department of Medicine & Epidemiology, School Veterinary Medicine, University of California, Davis, California, United States of America

[2] Centre for Veterinary Epidemiology and Risk Analysis (CVERA), UCD School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland

[3] Marine Institute, Rinville, Oranmore, County Galway, Ireland

[4] Marine Harvest Ireland, Rinmore, Letterkenny, County Donegal, Ireland

[5] California Animal Health and Food Safety Laboratories (CAHFS), Department of Medicine & Epidemiology, School Veterinary Medicine, University of California, Davis, California, United States of America

*Corresponding author

E-mail: tyatabe@ucdavis.edu (TY)

Abstract

Salmonid farming in Ireland is mostly organic, which implies limited disease treatment options, and its network of live fish movements possesses characteristics that would facilitate infection spread processes. This highlights the importance of biosecurity for preventing the introduction and spread of infectious agents. In this paper we characterized the biosecurity of the salmonid farms in Ireland using a survey, then developed a score for benchmarking the disease risk of salmonid farms. The usefulness and validity of this score - together with farm indegree, for predicting disease diversity, defined as the number of different diseases affecting a farm during a year, was assessed through generalized Poisson regression using Bayesian inference. Seawater salmon (SW salmon) farms had the highest biosecurity scores with a median (IQR) of 82.3 (5.4), followed by freshwater salmon (FW salmon), 75.2 (8.2), and freshwater trout (FW trout) farms 74.8 (4.5). For FW salmon and trout farms, the top ranked model (in terms of leave-one-out information criteria, looic) was the null model (looic = 46.1). For SW salmon farms, the best ranking model was the full model (looic = 33.3). Farms with a higher biosecurity score were associated with lower disease diversity, and farms with a high number of suppliers (indegree > 1) were associated with increased disease diversity. The effect of the interaction between these variables was also important, showing an antagonistic effect, which would indicate that biosecurity effectiveness is achieved through a broader perspective on the subject, which includes a minimization in the number of suppliers and hence in the possibilities for infection to enter a farm. The work presented here could be used to elaborate indicators of a farm's disease risk based on its biosecurity score and number of suppliers, to better inform risk-based disease surveillance and control strategies for private and official stakeholders.

Introduction

Commercial salmonid farming has been present in Ireland since 1979, being a significant contributor to the Irish economy, particularly along the western seaboard of the country. In Ireland, as in other countries where Atlantic salmon (*Salmo salar*) is produced, the production system is roughly divided into 3 main types of farms: broodstock farms, where eggs and milt are obtained from sexually mature fish, and are mixed to produce fertilized eggs, which are later grown on farm or transported to other farms to further develop; freshwater farms, most of which have a hatchery where fertilized eggs hatch and fish are kept until first feeding stage, to be moved within the farm to larger production tanks, where fish will grow until becoming smolts (roughly 70-100 grams, 10-15 months of age), the stage where fish are ready to transition into the ocean (although some companies further move the fish to net pens in freshwater lakes for the smoltification to occur there); finally, seawater farms to which smolts are transported, where they will grow until they have reached market size (about 4-5 kilos, 18 to 24 months of age) or until they become sexually mature, if they are selected to become the broodstock for the next production cycle. Atlantic salmon eggs and milt are sourced from local broodstock farms, but also imported into the country from other European countries [1].

The rainbow trout (*Oncorhynchus mykiss*) industry in Ireland is based on egg imports from within Europe and the USA [1], which are grown on freshwater farms until harvest, or sold to other freshwater farms for further growing, or to angler's clubs for recreational purposes. Farms that harvest fish could either process the fish on farm or send harvested fish to be processed in farms that have a processing plant.

This intricate movement of fish within and between the different types of salmonid farms in Ireland could be represented as a network, and hence it is amenable for network analysis. Our previous work has shown that the network of live salmonid fish movements in Ireland possesses characteristics that would facilitate infection spread processes, namely: a power-law degree distribution, short average path length and high clustering coefficients, when compared to random networks of the same order and volume [2]. This network structure determines the presence of farms that could potentially act as super-spreaders or super-receivers of disease, with few intermediaries of fish movement between farms, where infectious agents could easily spread, provided no effective barriers are placed within these farms. Additionally, all of the Irish salmon farming is certified organic, and hence farmers have a very limited therapeutic arsenal at their disposal to treat disease [3].

The above-mentioned network structure and predominant organic nature of Ireland's salmonid farming industry highlights the importance of biosecurity, as a means of preventing the introduction and or spread of infectious agents to farms within the Irish salmonid farming industry. In aquaculture, biosecurity has been defined as the sum of all procedures in place to protect living organisms from contracting, carrying, and spreading infectious agents and other non-desirable health conditions [4]. Effective biosecurity strategies provide protection to both farmed and wild aquatic animal populations, by minimizing the risk of introducing pathogens, and minimizing the consequences or further spread if the pathogen was introduced [5].

The objectives of this paper are three-fold: first, to characterize the biosecurity of the salmonid farms in Ireland, both in the freshwater and seawater environments using a survey based approach; second, based on the survey's results, to develop a score for benchmarking the biosecurity levels of salmonid farms in Ireland; and third, test the usefulness and validity of this score for predicting a farm's disease risk while, based on the network characteristics previously mentioned, accounting for the effect of farm centrality measures.

Methods

Biosecurity characterization

For the first objective of this research, a biosecurity survey was designed, based on the Scottish Code of Good Practice [6] and the recommendations for biosecurity of the Southern Regional Aquaculture Center [7-9]. The survey consisted mostly of closed questions and was divided into the following areas: farm stocking and characteristics, predator control, cleaner fish (seawater salmon farms only), disease prevention and control, divers and diving equipment (seawater farms only), handling of mortalities, feed and farm management, harvesting (seawater salmon and freshwater trout farms only), coordinated bay management and sea lice/amoebic gill disease monitoring and control (seawater farms only), fish welfare and care, management of people, and biosecurity program and records. In all, there were 75 questions for Atlantic salmon freshwater farms, 108 questions for Atlantic salmon seawater farms, 89 questions for Atlantic salmon lake farms, and 80 questions for freshwater trout farms. These surveys were critically reviewed by members of the industry, regulatory bodies and academia, and later piloted through administration at a research Atlantic salmon hatchery and a marine Atlantic salmon farm. Administration of the survey was conducted in person by the first author to salmonid farm managers during September and October of 2015. At the Atlantic salmon farms, the interview was conducted at 18 out of 20 active seawater farms, 14 out of 21 active freshwater farms, and 1 out of 3 active lake farms. At the trout farms, interviews were conducted at 8 out of 9 active freshwater farms. The duration of the survey administration ranged between 1 and 2 hours. The surveys are available upon request to the corresponding author.

Biosecurity scoring system

For the second objective of this paper, a score for each farm was calculated based on the manager's response to each of the questionnaire's closed questions. Each question had a maximum attainable score (i.e. the response(s) that was deemed, based on the above mentioned references, to make a farm least vulnerable to disease introduction and spread), from which points were discounted if other, less optimal options were being carried out at the farm per the manager. For example, for the question "Does the farm receive fish from other seawater farms on a typical production cycle?" the maximum attainable score was 1 point if the answer to this question was 'no'. Another example for seawater farms is the question "How do you make sure fish that arrive at your farm are in good condition?" A farm would get a maximum score of 6 points for this question if each of the following were requested by the farm manager: health certificates, diagnostic test results, a sanitary history, and smoltification test results from its suppliers, together with inspection prior to purchase and upon arrival. One point would be deducted from this maximum score for each of these measures that were not reported by the farm manager.

Each question had two quantities associated with it: the score obtained by the farm, and the maximum potential score attainable. Therefore, the computation of the overall biosecurity score for a farm i was based on the following formula

$$\frac{\sum obtained\ points_i}{\sum potential\ points_i} x100 = biosecurity\ score_i \tag{1}$$

Where the score is scaled by 100 to make it a percentage of the "maximum attainable biosecurity". Open questions were excluded as they were few, mostly descriptive, and with varying levels of completion or detail by interviewed managers, and hence not readily amenable for a scoring system of this kind.

Disease diversity models

The only open question that was kept, although not used in the scoring of a farm, was the list of different diseases that occurred in the farm in the 12 months prior to the survey (termed 'disease diversity' subsequently). The results from this question were modeled as originating from a generalized (or Lagrangian) Poisson process, as defined by Consul and Jain (10), with a probability mass function given by

$$P(N = n) = p_n(\theta, \lambda, m) = \begin{cases} \theta(\theta + n\lambda)^{n-1\frac{\exp(-\theta-n\lambda)}{n}} & for\ n = 0, 1, 2, \dots, m \\ 0 & for\ n > m\ when\ \lambda < 0 \end{cases} \tag{2}$$

Where $\theta > 0$, $\max(-1, -{}^{\theta}/_{m}) \leq \lambda \leq 1$, and m taken equal to the largest possible integer such that $\theta + m\lambda > 0$ when λ is negative [11]. The expectation and variance of this distribution are given by

$$E(N) = \frac{\theta}{1-\lambda} \ , \ Var(N) = \frac{\theta}{(1-\lambda)^3} \tag{3}$$

From this it can be seen that this distribution allows to model data which shows either over-dispersion (for which $\lambda > 0$) or under-dispersion (for which $\lambda < 0$), and that the generalized Poisson distribution (GPD) reduces to the Poisson distribution when $\lambda = 0$ [12].

From an interpretative point of view, this distribution extends the Poisson distribution by its ability to describe situations where the probability of occurrence of a single event does not remain constant (as in a Poisson process), but is affected by previous occurrences [13]. The generalized Poisson distribution has been found to accurately describe phenomena as diverse as the observed number of industrial accidents and injuries, where a learning effect may be present; the spatial distribution of insects, where initial occupation of a spot by a member of the species has an influence on the attractiveness of the spot to other members of the species; and the number of units of different commodities purchased by consumers, where the current sales have an impact on the level of subsequent sales through repeat purchases [14]. Similarly, the number of different diseases affecting a farm could be thought of as arising in a similar manner, where the occurrence of an outbreak of disease is influenced by other disease events that have taken place before in the farm. This effect could arise from synergy or antagonism between agents, or from a learning curve or changes in the level of awareness in farm personnel.

To assess the effect of in-farm biosecurity, measured as a biosecurity score, and farm centrality, indegree specifically, a generalized Poisson regression model [12] was fit to the data of the form

$$\log\big(E(N|biosec, indegree)\big) = \log\left(\frac{\theta}{1-\lambda}\right) = \beta_0 + \beta_1 biosecurity + \beta_2 indegree + \beta_3 interaction \tag{4}$$

Where $E(N|biosec, indegree)$ is the expected disease diversity given the farm biosecurity score and indegree, $\theta = \exp(\beta_0 + \beta_1 biosecurity + \beta_2 indegree + \beta_3 interaction)\,(1 - \lambda)$, β_0 is the intercept, β_1 is the regression coefficient for the farm's biosecurity score, β_2 is the regression coefficient for the farm's indegree, and β_3 is the regression coefficient for the interaction between biosecurity and indegree. The values of λ and θ were restricted as described above in the support of the distribution.

Two of these models were fitted: one for seawater farms and another one for freshwater farms, which included salmon hatcheries, lake farms, and freshwater trout farms. These will be referred to as SW and FW models from now on. The largest possible value, m, was set equal to 3 for the SW model and 2 for the FW model, as this was the maximum number of diseases reported in these types of farms during the survey.

Model fitting was carried out in a Bayesian framework, with priors

$$\beta_0, \beta_1, \beta_2, \beta_3 \sim Normal(0, 0.5) \tag{5}$$

$$\lambda \sim Normal(0, 1) \tag{6}$$

For these models, indegree was transformed into a binary variable, where the value 0 was for a farm with indegree = 1 (i.e. only one fish supplier), and 1 for a farm with indegree > 1 (i.e. 2 or more fish suppliers) during the year prior to the survey (2014). For 3 farms that did not report any fish movements during 2014, data from 2013 was used. We will refer to this dichotomized variable as number of suppliers from now on. The biosecurity score was scaled to have a mean of 0 and standard deviation of 0.5, and the number of suppliers variable was centered at its mean. These transformations allow for the interpretation of the regression coefficients to be more transparent, by making them directly comparable from a parameter estimates table [15], although plotting is a much clearer mean of understanding the predictor variables' effects and interactions, and hence the main approach used here.

For both SW and FW models 5 models were fit: i) the full model with two main effects (biosecurity score and number of suppliers) and their interaction, ii) main effects (no interaction) model, iii) a model with biosecurity alone, iv) a model with number of suppliers alone, and v) a null model with no predictors in it. Comparisons between these models were done using leave-one-out cross-validation criterion (looic) [16].

Each model was initially fitted using 4 chains of 4,000 iterations with a warm-up of 2,000 iterations for assessing model convergence, after which an individual chain of 16,000 iterations with warm-up of 8,000 iterations was used for inference for each model. Model convergence diagnostics included visual checking of trace plots, to visually evaluate stationarity and mixing of the chains, Gelman-Rubin convergence diagnostic, $\hat{R}$,and the number of effective samples [17, 18].

The model was fit using Stan's Hamiltonian Monte Carlo sampling [19] in the R statistical environment [20], using the Rstan package [21]. Leave-one-out cross-validations were done using the loo package in R [22].

Finally, posterior prediction checks were used to evaluate if simulated samples from the model matched the original data. This was done visually through histograms and comparing the mean and variance of simulated and observed data. For generating random samples from a generalized Poisson distribution, the package RMKdiscrete [23] was used.

Results

Biosecurity characterization

In total, 21 farm managers completed the survey. This corresponds to 35 farms (66% of active salmonid farms in Ireland during the survey period), as some managers were in charge of 2 or more farms. Of these farms, 27 (77%) were Atlantic salmon and 8 (23%) were freshwater trout farms. Of the salmon farms, 18 were seawater farms, 8 were freshwater farms, and 1 was a freshwater lake farm. Most of the trout farms were inland farms where fish were produced for grow out in other inland farms or for consumption, except for a small scale farm which raised trout for repopulation purposes. One of the seawater farms was fallowed since 2013, so it was not further considered for the disease diversity models. Biosecurity survey results are included in supporting information S1, S2, and S3 for Atlantic salmon seawater, Atlantic salmon freshwater, and freshwater trout farms, respectively.

Biosecurity score, indegree, and disease diversity

Regarding the biosecurity score, in general seawater salmon farms had the largest values, followed by freshwater salmon farms and trout farms, the latter two being very similar. Indegree was also highest among seawater farms, followed by freshwater salmon farms, with freshwater trout farms being the lowest. With respect to disease diversity, the largest values where for freshwater trout and seawater salmon farms, with freshwater salmon farms showing the least diversity (Table 1).

Table 1. Descriptive statistics for the biosecurity score, indegree and disease diversity for the surveyed farms

Type (N)	Biosecurity score			Indegree			Disease diversity		
	1st quartile	Median	3rd quartile	1st quartile	Median	3rd quartile	1st quartile	Median	3rd quartile
SW salmon (18*)	82.3	84.1	87.7	1.0	2.0	2.0	1.0	1.0	2.0
FW salmon (8)	70.3	75.2	78.5	0.8	1.0	2.0	0.0	0.0	0.3
FW lake salmon (1)	NA	94.4	NA	NA	1.0	NA	NA	0.0	NA
FW trout (8)	71.8	74.8	76.3	0.0	0.5	1.0	1.0	1.0	2.0

SW: seawater; FW: freshwater; * indegree and disease diversity scores based only on 17 seawater farms

The most common disease affecting seawater salmon farms was pancreas disease (PD) caused by the salmonid *alphavirus*, 14 farms, followed by amoebic gill disease (AGD) caused by *Neoparamoeba* *perurans*, 10 farms, with one farm reporting the occurrence of an infectious pancreatic necrosis (IPN) outbreak. Only one farm reported that it had not experienced any disease outbreak in the preceding year, 8 farms reported experiencing one disease (either PD or AGD), 7 farms reported experiencing 2 diseases (both PD and AGD), and one farm reported experiencing all three diseases in the previous year. For freshwater salmon farms only 2 farms reported to have experienced diseases in the preceding year, 1 farm reported an outbreak of furunculosis (*Aeromona salmonicida subsp. salmonicida*), and the other one reported three diseases: *Ichthyobodo sp*., and unspecified gill and fungal infections. All but one freshwater trout farms reported experiencing disease in the preceding year: 2 farms reported bacterial gill disease (BGD) caused by *Flavobacterium sp*., one farm reported rainbow FW trout fry syndrome (RTFS) caused by *Flavobacterium psychrophilum*, and 4 farms reported both RTFS and *Ichthyobodo sp*.

Disease diversity models

For the disease diversity models, model comparisons are presented in Table 1 for the 2 settings (seawater and freshwater). For the freshwater farms the best ranking model was the null model (leave-one-out cross-validation information criteria, looic = 46.1), indicating that none of the predictors shows a meaningful association with the number of diseases affecting a freshwater farm, although the difference with the model with only biosecurity score (looic = 46.2) was negligible. For the seawater farms the best ranking model was the full model (looic = 33.3), which included the biosecurity score, number of suppliers, and their interaction.

Table 1. Model comparisons for disease diversity models for seawater and freshwater farms using leave-one-out information criterion (looic)

Seawater farms				
Model	**looic**	**SE looic**	**p-loo**	**SE p-loo**
Full model	33.3	3.1	1.9	0.7

Biosecurity score	39.3	4.1	1.8	0.8
Null model	39.9	5.3	1.5	0.7
Main effects	40.1	4.2	2.3	0.8
No of suppliers	41.4	5.4	2.1	0.9
Freshwater farms				
Null model	46.1	5.4	1.5	0.3
Score	46.2	5.3	1.7	0.3
No of suppliers	46.8	5.4	1.8	0.3
Main effect	47.0	5.3	2.0	0.3
Full model	47.3	5.4	2.2	0.3

p-loo: number of effective parameters

For the latter model, both the biosecurity score and number of suppliers seemed to be important predictors (most of the probability mass was away from the null value of zero), with farms having a higher biosecurity score being associated with a lower disease diversity affecting the farm, and farms with a high number of suppliers (more than one supplier or indegree greater than 1) being associated with an increased disease diversity. The effect of the interaction between both variables also seemed to be important, indicating these two variables modulate each other's effect (Table 2 and Figs 1 and 2).

Table 2. Parameter estimate's posterior distribution of disease diversity's top ranked seawater model and an equivalent model for freshwater farms

Seawater farms				
Parameter	**mean**	**SD**	**2.5%**	**97.5%**
Intercept	0.24	0.12	0.01	0.49
Biosecurity score	-0.27	0.21	-0.68	0.16
No of suppliers	0.25	0.23	-0.22	0.68
Interaction	0.70	0.33	0.03	1.35
lambda	-0.71	0.20	-0.99	-0.25
Freshwater farms				
Intercept	-0.10	0.25	-0.61	0.40
Biosecurity score	-0.23	0.37	-0.95	0.50
No of suppliers	-0.07	0.40	-0.87	0.72
Interaction	0.02	0.49	-0.95	0.98
lambda	0.10	0.20	-0.27	0.51

Fig 1. Probability distribution of the parameters of the top ranked seawater farm model (left), and equivalent model for freshwater farms (right). Black dot: median, red thick line: 66% PI, black thin line: 95% PI.

Fig 2. Model's estimate of the rate of different diseases affecting a seawater farm during a year for farms with high (indegree > 1) and low (indegree = 1) number of suppliers. Farm with a high number of suppliers: median (solid blue line), 66% PI (purple shaded area), and 95% PI (pink shaded area); Farm with a low number of suppliers**:** median (dashed grey line), 66% PI (grey shaded area), and 95% PI (green shaded area). Observed data for farms with a high number of suppliers (filled points) and low a number of suppliers (hollow points). A jitter was added to the observed data points to avoid superimposition in the plot.

For the effect of biosecurity, although crossing the null value of zero (95% PI of -0.68, 0.16), most of the posterior distribution lies below this value, with a median of -0.27. Similarly, the posterior distribution of the effect of the number of suppliers is mostly above zero, with a median of 0.25 (95% PI -0.22, 0.68). In the case of the interaction term between these two variables, it has a median of 0.70 (95% PI 0.03, 1.35). Lambda, the dispersion parameter of the distribution had a mean posterior value of -0.71 (95% PI -0.99, -0.25), indicating that the disease diversity a seawater farm experiences could be thought of as an under-dispersed Poisson process.

These results indicate that the variables included in the model are only explanatory for seawater salmon farms. For these farms, the effects of the biosecurity score and number of suppliers are better appreciated in Fig 2. This figure shows that the effect of biosecurity is modulated by the number of suppliers of the farm. If the number of suppliers is low (i.e. only one fish supplier), biosecurity seems to have a protective effect, reducing the expected number of diseases affecting a seawater farm, as shown by the black dashed line on the plot. On the other hand, if number of suppliers is high (i.e. more than one supplier of fish) the effect of an increasing biosecurity seems to be negligible, as shown by the mostly flat horizontal line in the plot.

Fig 3 further explores this antagonistic effect, showing the estimated difference in the rate of different diseases affecting farm, varying one of the variables (biosecurity score or number of suppliers) while keeping the other constant. Specifically, the upper left plot shows the difference in rate for a farm with high number of suppliers (indegree > 1) versus a farm with a low number of suppliers (indegree = 1), when both farms have the same low biosecurity score: here we see that difference in the rate is virtually negligible, with a median of -0.55 (95% PI of -1.64, 0.62). On the other hand, the upper right plot shows the same comparison for farms with high biosecurity, with a median difference in the rate of 0.71 (95% PI of -0.01, 1.39). Regarding the effect of differences in biosecurity, the lower left plot shows that for a farm with high number of suppliers, the rate difference between a farm with low vs a farm with high biosecurity is almost inexistent, with median of -0.04 (95% PI of -1.05, 1.29), while for a farm with low number of suppliers (lower right plot), the difference is substantial, with a median of 1.22 (95% PI of 0.19, 2.32).

Fig 3. Posterior density of the difference in the rate of different diseases affecting a seawater salmon farm during a year. a) high (indegree > 1) vs low (indegree = 1) number of suppliers for a low (79.5) biosecurity score, b) high vs low number of suppliers for a high (88.6) biosecurity score, c) low vs high biosecurity for a high number of suppliers, d) low vs high biosecurity for a low number of suppliers. Shaded area: 95% PI of the difference; dashed vertical line: median difference

The distribution of 136,000 simulated counts of the number of different diseases affecting a seawater salmon farm during a year, with a mean and variance of 1.40 and 0.65, is very similar to the one presented by the data, which had a mean and variance of 1.47 and 0.51 (Fig 4). Only 0.07% of the simulated counts were greater than 3 (not shown in the plot). As a means of comparison, for the more typical Poisson regression (which assumes lambda equals 0), with the same predictors, the mean and variance were 1.46 and 1.84, respectively, with 8.1% of simulated counts greater than 3, and a 27.4% of simulated counts of zero (as opposed to only 11.8% of the generalized Poisson model and 5.9% from the observed data). This indicates that the under-dispersed generalized Poisson seems to be a more suitable model than the Poisson as the originating process for the data at hand.

Fig 4. Histogram of simulated (left) and observed (right) number of diseases affecting a seawater salmon farm during a year. Simulations that resulted in more than 3 different diseases (0.7% of the 136,000 simulated counts) were omitted.

Discussion

In this study we have characterized the biosecurity of salmonid farms in Ireland using a survey based approach, developed a biosecurity score to benchmark farms propensity to disease outbreaks, described the differences in number of suppliers between types of farms, the disease diversity (this being the number of different diseases that occurred in the 12 months prior to the survey) affecting the farms, and identified the interplay between biosecurity, number of suppliers and disease diversity for seawater Atlantic salmon farms.

Of the salmonid farms in Ireland, results would indicate that seawater salmon farms would have the highest biosecurity levels, with freshwater salmon and trout farms having the lowest values. However, comparisons between different production settings could be misleading, as the instruments of measurement (surveys) were different for each type of farm, being this in turn a reflection of the different nature of the environments on which these production phases take place. Because of this, results should not be interpreted as indication of increased disease risk in freshwater compared to seawater farms but rather of the potential for improvements in biosecurity in these premises. Therefore, the biosecurity score developed here should be considered a tool of benchmarking farms within the same environment, rather than comparing biosecurity between different production settings.

At freshwater salmon farms, fish are less exposed to disease causing agents due to a higher ability to control the environment, whereas higher environmental exposure at seawater salmon farms is the norm. This could explain the lower disease diversity found in freshwater salmon farms. The higher disease diversity in freshwater trout farms (which had a similar score to freshwater salmon farms) could be related to the seemingly less controlled environmental conditions of these type of farms in Ireland, with several farms having earth pond rearing systems, with very little, if any, possibility to do a thorough cleaning between generations of fish, a constant mixing of fish generations, less control over water quality parameters, and less efficient mortality removal and predator control practices.

The absence of an apparent effect of the biosecurity score on disease diversity for freshwater salmon and trout farms by no means undermines the importance of biosecurity on these farms. In fact, in the freshwater model, most of the probability mass for the biosecurity score effect lies below zero (Fig 1). Perhaps further data will provide more support for the protective effect of the biosecurity score for these types of farms. On the other hand, the effect of number of suppliers and the interaction between this and biosecurity does not seem to be supported by available data for these types of farms.

The virtual absence of contacts for freshwater trout farms (median and maximum indegree of 0.5 and 1, respectively) is because most of the freshwater trout farms in Ireland import their eggs from abroad, which was not taken into account in this study, with very little movement between farms. For the freshwater salmon farms, there were only two farms that did not receive fish from other farms within Ireland: one farm imported all of its stock (eggs) from abroad, and the other one was a research facility that participated in a population enhancement program, capturing wild broodstock and releasing juveniles to the environment. This in turn could explain the absence of any meaningful effect of number of suppliers on disease diversity, as the freshwater farms with higher variability in number of suppliers (salmon freshwater farms) were the farms with lowest variability in disease diversity, whereas freshwater trout farms, where variability for disease diversity was high, had very little variability in their domestic number of suppliers.

Model results indicate that biosecurity is important in determining the susceptibility of salmon seawater farms to different diseases, expressed as disease diversity, with farms with higher biosecurity presenting minor diversity in their disease burden. As far as the authors know, the effect of biosecurity on a farm's disease diversity has never been studied. Also, seawater farms with low number of suppliers (indegree < 2) would also have lower disease diversity during the production cycle. The effect of increased fish movement on disease risk has been explored before for specific diseases such as ISA [24], and the effect of multiple smolt suppliers has been also associated with an increased risk of specific diseases, such as IPN [25]. Nevertheless, to the best of our knowledge, this is the first study that evaluates the putative effect of the number of suppliers on disease diversity in an animal production setting.

Regarding the specific diseases that comprised the disease diversity of a seawater farm, the most common one was pancreas disease. This is a viral infectious disease caused by the salmonid alphavirus pancreas disease virus (PDV) [26, 27]. The second most common disease was amoebic gill disease (AGD) whose causative agent, *Neoparamoeba perurans*, is considered an environmental free living protozoan [28, 29]. Finally, only one farm reported an outbreak of infectious pancreatic necrosis, which is caused by the IPN virus [30].

The effect of the biosecurity score in the incidence of AGD (a facultative pathogen) could be a reflection of the association between biosecurity and good husbandry practices, which, together with other environmental and host factors, determine whether or not a pathogen causes an overt disease [31]. Similarly, the effect of the number of suppliers to a seawater salmon farm on AGD could be considered an indicator of fish stress (for both new and resident fish) associated with stocking new fish in the farm several times, which in turn would affect the ability of fish to resist disease [32]. For the other two diseases that comprised the disease diversity of seawater farms, the viral infectious diseases PD and IPN, to the above mentioned underlying factors, it could also be added the higher vulnerability to disease introduction and spread entailed by a lower biosecurity score.

The estimated antagonistic effect between the protection provided by biosecurity, measured as a biosecurity score, and the number of suppliers seems very meaningful. The results here indicate that potential benefits of biosecurity in terms of reducing a farm's disease diversity are "diluted" by a high number of suppliers (Fig 4c), and is only beneficial for seawater farms with a low number of suppliers (Fig 4d). Similarly, the positive effect of a low number of suppliers (i.e. no more than one supplier) in reducing disease diversity is only manifest when biosecurity is high (Fig 4b). This antagonistic effect suggests that maximum biosecurity effectiveness is only achieved through a broader perspective on the subject, with biosecurity including efforts to minimize the number of suppliers and hence the possibilities for introduction of infection into a farm. The multiple supplier effect in reducing the effectiveness of biosecurity could be related to increased instances of fish stress (one for each stocking, potentially many for each supplier), and to the multiple chances of introducing infected fish populations to the farm.

Regarding the modeling approach used, a generalized Poisson distribution was chosen to capture the potential lack of independence between different disease events in a farm, due to interaction between different diseases, and in-farm learning curve/change in disease awareness. It also allowed us to flexibly model a count process as the one presented here, allowing for counts that are either under-, equi-, or over-dispersed[12].

Possible limitations of this study include its cross-sectional nature, where the survey was administered during a 2 month span, with questions regarding both exposure (biosecurity) and outcome (diseases affecting the farm in the year prior to the survey). This could have led to recall bias both for exposure and outcome, and ensuing misclassification. In spite of this, the authors consider that most misclassification, if present, was non-differential, as the part of the survey dealing with the outcomes was located in the middle part of the survey, the interviewed managers did not seem to show any negative attitude when prompted about the diseases affecting the farm, and confidentiality was assured at all times. Nevertheless, access to farm production records in the country should be considered in future research, including attributed mortality causes.

Another issue relates to the weight assigned to each question in the survey for estimating the biosecurity score. In the current study, all questions had the same weight. Nonetheless, it is possible that different weighting schemes would produce different results, increasing or decreasing the measure of effect. An attempt was made to identify the most important questions of the survey through matrix factorization, specifically principal component analysis, for the score of seawater salmon farms. Using this approach, the burden of questions was reduced from 108 to 55, and when fitted in a model equivalent to the one using the full data set, we were able to establish a similar interplay between biosecurity score, number of suppliers and, disease diversity, although the magnitude of the effect estimated with the score based on the full survey was much higher for both the score and the interaction with number of suppliers (results not shown).

Finally, the work presented here could be used to elaborate indicators of a farm's risk of disease occurrence based on biosecurity scoring system and its indegree, to inform risk-based disease surveillance and control activities at both the private and official sectors. Future work will include the analysis of in-farm mortality records, to evaluate if the associations found here are valid.

Acknowledgements

This research is part of a multi-institutional and multi-national project to characterize the risk of introduction and spread of infectious agents into local salmon farming industries in the context of globalization, whose partnering institutions are: Marine Harvest Ireland, the Fish Health Unit at the Marine Institute, the Centre for Veterinary Epidemiology and Risk Assessment (CVERA) at University College Dublin, and the Center for Animal Disease Modeling and Surveillance (CADMS) at the University of California, Davis. Authors would like to acknowledge to the Irish salmonid farmers, managers and staff who took part in the survey process.

References

1. Anon. Domestic live fish movement dataset: 2014. Marine Institute; 2015.
2. Yatabe T, More SJ, Geoghegan F, McManus C, Hill AE, Martínez-López B. Characterization of the live salmonid movement network in Ireland: Implications for disease prevention and control. Preventive Veterinary Medicine. 2015;122(1–2):195-204.
3. BIM. Annual Aquaculture Survey 2015. In: Department of Agriculture FatM, editor. Dublin, Ireland: Bord Iascaigh Mhara, Irish Sea Fisheries Board; 2016. p. 16.
4. Pruder GD. Biosecurity: application in aquaculture. Aquacultural Engineering. 2004;32(1):3-10.
5. Oidtmann BC, Thrush MA, Denham KL, Peeler EJ. International and national biosecurity strategies in aquatic animal health. Aquaculture. 2011;320(1–2):22-33.
6. A code of good practice for Scottish finfish aquaculture, (2006).
7. Yanong RP. Biosecurity in aquaculture, part 2. Recirculating aquaculture systems Southern Regional Aquaculture Center, Publication. 2012;4708.
8. Yanong RP. Biosecurity in aquaculture, Part 3: Ponds. SRAC Publication-Southern Regional Aquaculture Center. 2013(4712).
9. Yanong RP, Erlacher-Reid C. Biosecurity in aquaculture, part 1: an overview. USDA Southern Regional Aquaculture Center. 2012;4707.
10. Consul PC, Jain GC. A Generalization of the Poisson Distribution. Technometrics. 1973;15(4):791-9.
11. Scollnik DP. On the analysis of the truncated generalized Poisson distribution using a Bayesian method. Astin Bulletin. 1998;28(01):135-52.
12. Consul P, Famoye F. Generalized Poisson regression model. Communications in Statistics-Theory and Methods. 1992;21(1):89-109.
13. Tuenter HJH. On the generalized Poisson distribution. Statistica Neerlandica. 2000;54(3):374-6.
14. Consul PC. Generalized Poisson distributions: properties and applications. New York: Marcel Dekker Inc.; 1989.
15. Gelman A. Scaling regression inputs by dividing by two standard deviations. Statistics in medicine. 2008;27(15):2865-73.
16. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing. 2016:1-20.

17. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis. 3 ed: Chapman & Hall/CRC Boca Raton, FL, USA; 2014.
18. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Statistical science. 1992:457-72.
19. Stan Development Team. Stan Modeling Language Users Guide and Reference Manual. Version 2.14.0 ed2016.
20. R Development Core Team. R: A language and environment for statistical computing. In: Computing RFfS, editor. Vienna, Austria2017.
21. Stan Development Team. RStan: the R interface to Stan. 2.14.1 ed2016.
22. Vehtari A, Gelman A, Gabry J. loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. 2016. p. R package version 1.0.
23. Kirkpatrick R. RMKdiscrete (Version 0.1). Software and documentation available at http://cran r-project org/web/packages/RMKdiscrete. 2014.
24. Murray AG, Smith RJ, Stagg RM. Shipping and the spread of infectious salmon anemia in Scottish aquaculture. Emerg Infect Dis. 2002;8(1):1-5.
25. Jarp J, Gjevre AG, Olsen AB, Bruheim T. Risk factors for furunculosis, infectious pancreatic necrosis and mortality in post-smolt of Atlantic salmon, Salmo solar L. J Fish Dis. 1995;18(1):67-78.
26. Rodger H, Mitchell S. Epidemiological observations of pancreas disease of farmed Atlantic salmon, Salmo salar L., in Ireland. J Fish Dis. 2007;30(3):157-67.
27. Graham DA, Fringuelli E, Wilson C, Rowley HM, Brown A, Rodger H, et al. Prospective longitudinal studies of salmonid alphavirus infections on two Atlantic salmon farms in Ireland; evidence for viral persistence. Journal of Fish Diseases. 2010;33(2):123-35.
28. Downes JK, Henshilwood K, Collins EM, Ryan A, O'Connor I, Rodger HD, et al. A longitudinal study of amoebic gill disease on a marine Atlantic salmon farm utilising a real-time PCR assay for the detection of Neoparamoeba perurans. Aquaculture Environment Interactions. 2015;7(3):239-51.
29. Mitchell SO, Rodger HD. A review of infectious gill disease in marine salmonid fish. Journal of Fish Diseases. 2011;34(6):411-32.
30. Roberts RJ, Pearson MD. Infectious pancreatic necrosis in Atlantic salmon, Salmo salar L. J Fish Dis. 2005;28(7):383-90.
31. Plumb JA, Hanson LA. Health maintenance and principal microbial diseases of cultured fishes: John Wiley & Sons; 2011.
32. Bernoth E-M, Ellis AE, Midtlyng PJ, Olivier G, Smith P. Furunculosis: multidisciplinary fish disease research: Academic Press; 1997.

Supporting information

S1 File. Marine farms biosecurity survey results

S2 File. Freshwater farms biosecurity survey results

S3 File. FW trout farms biosecurity survey results

Figure1

Click here to download Figure Fig1.tiff

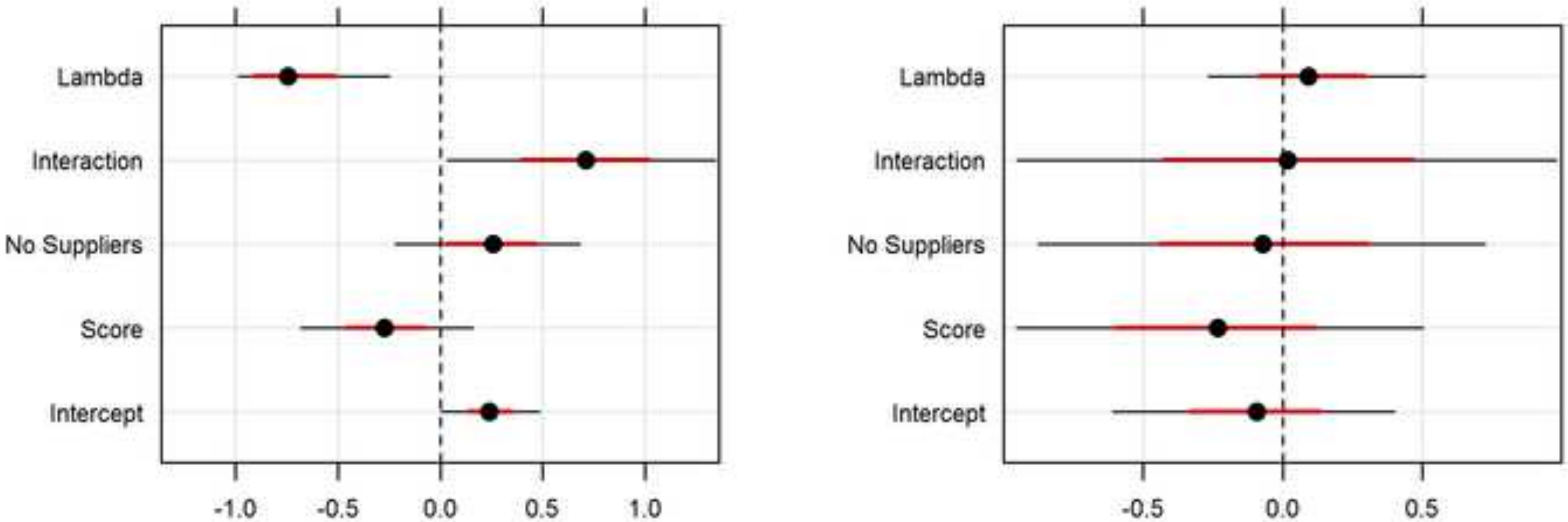


Figure2

Click here to download Figure Fig2.tiff

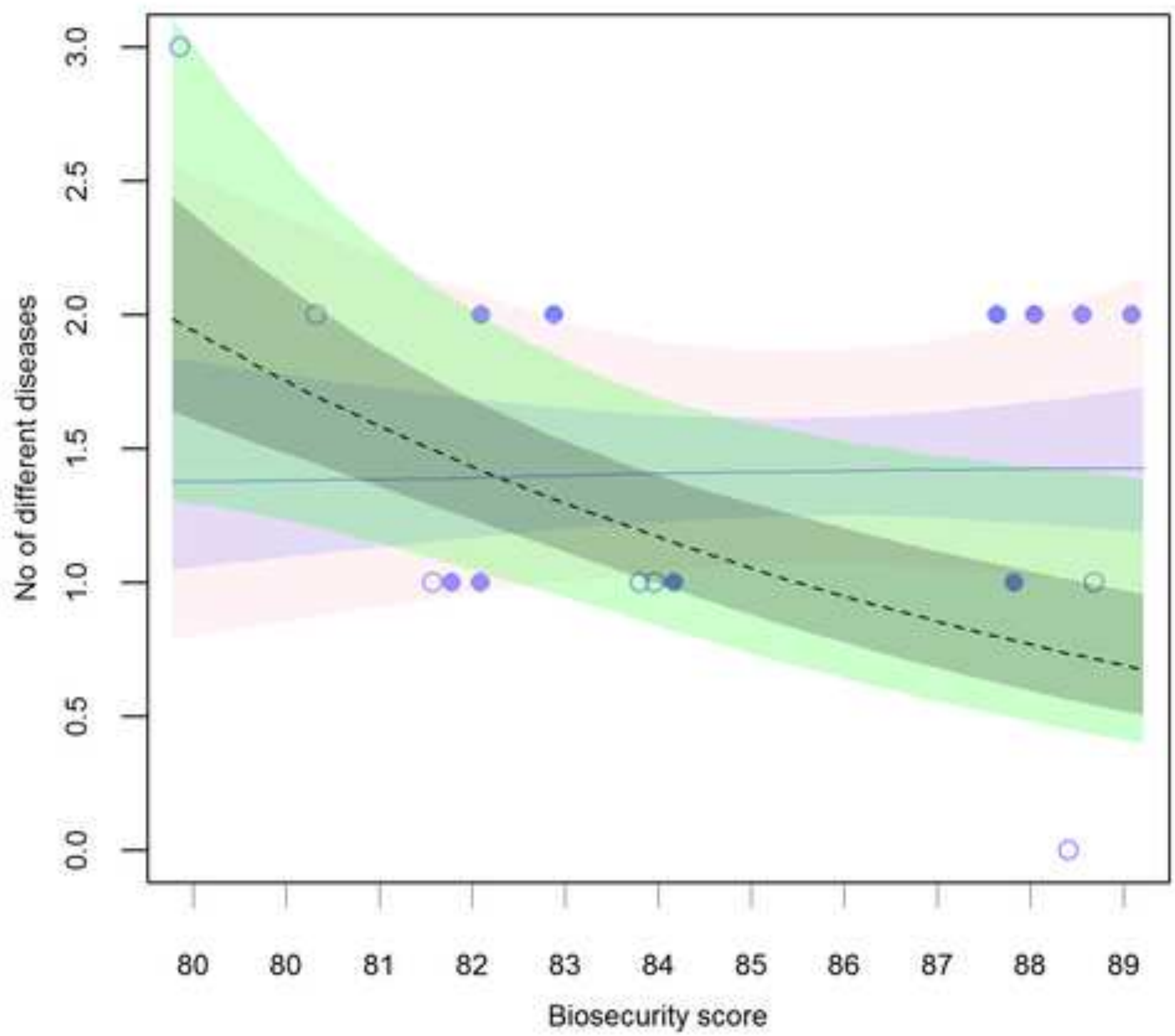


Figure3

Click here to download Figure Fig3.tiff

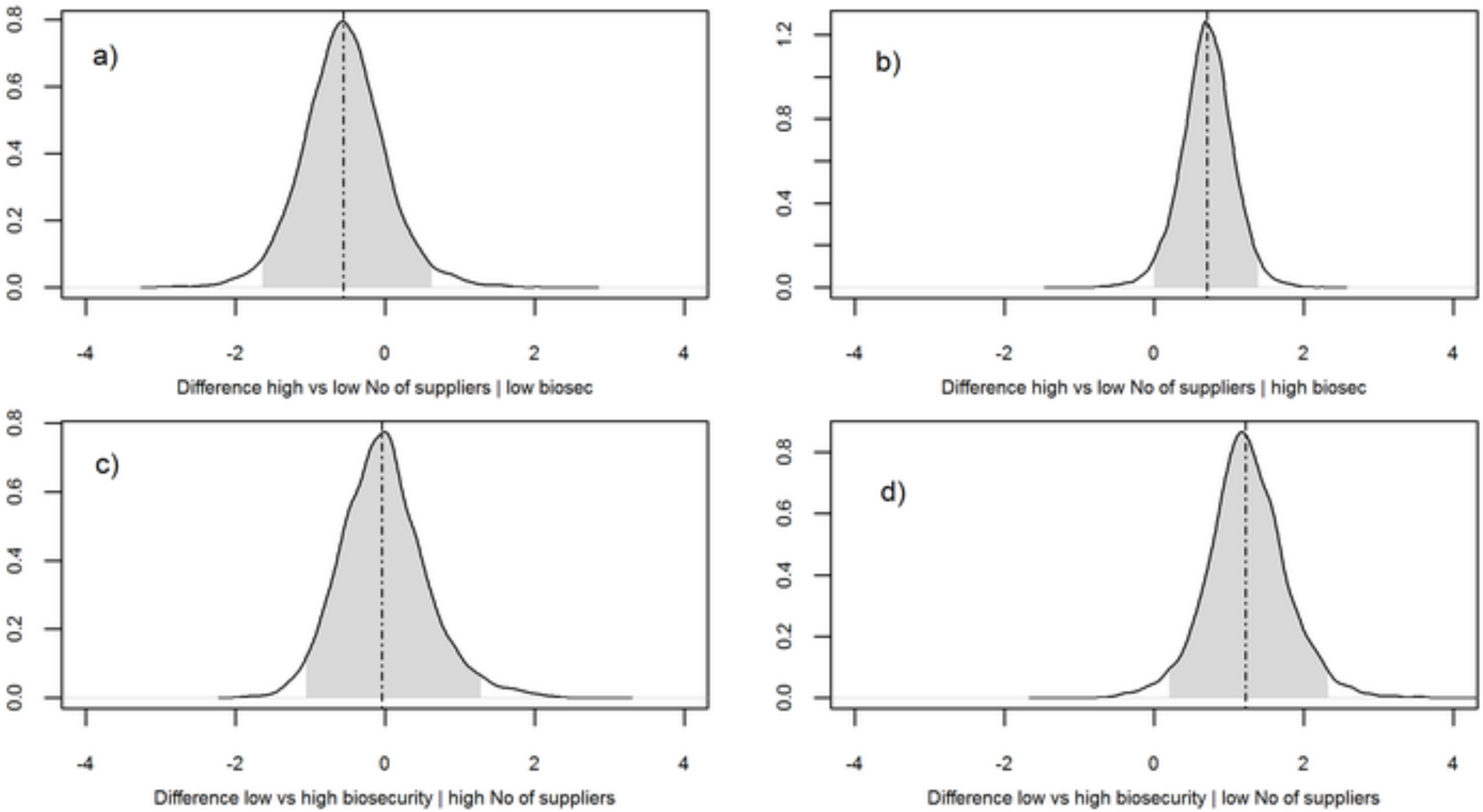
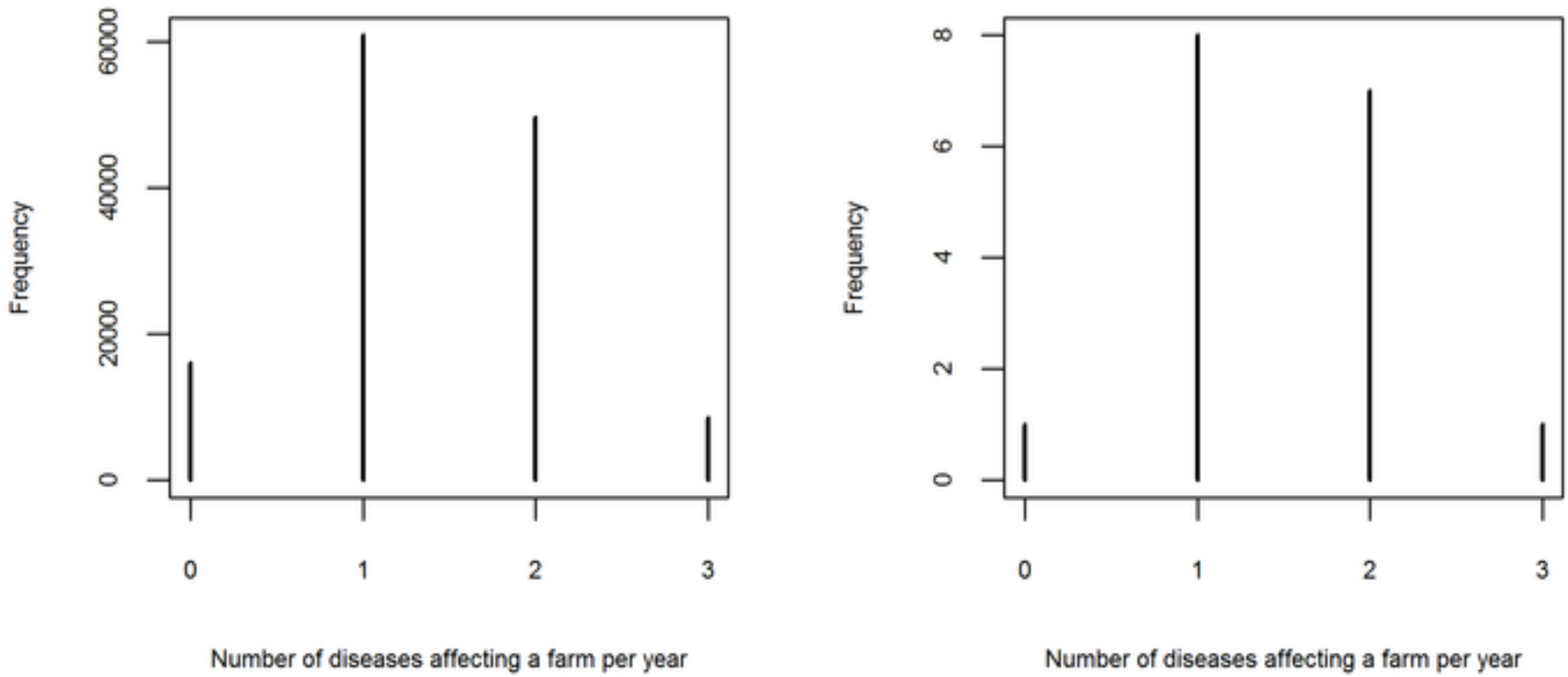


Figure4

Click here to download Figure Fig4.tiff



Supporting Information2

Click here to access/download

Supporting Information

S2_File.docx