

Bovine tuberculosis (*Mycobacterium bovis*) outbreak duration in cattle herds in Ireland: a retrospective observational study

Andrew W. Byrne^{1*}, Damien Barrett¹, Philip Breslin², Jamie M. Madden³, James O’Keeffe², and Eoin Ryan^{2*}

¹ One-Health Scientific Support Unit, Surveillance, Animal by-products, and TSEs (SAT) Division, Department of Agriculture, Food and the Marine, Agriculture House, Dublin 2, Ireland

² Ruminant Animal Health Division, Department of Agriculture, Food and the Marine, Backweston, Co. Dublin, Ireland

³ Centre for Veterinary Epidemiology and Risk Analysis (CVERA), School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland

* Correspondence: ER eoin.ryan@agriculture.gov.ie; AWB ecologicalepidemiology@gmail.com

Bovine tuberculosis (bTB) outbreaks, caused by *Mycobacterium bovis* infection, are a costly animal health challenge. Understanding factors associated with the duration of outbreaks, known as breakdowns, could lead to better disease management policy development. We undertook a retrospective observational study (2012-2018) and employed Finite Mixture Models (FMM) to model the outcome parameter, and to investigate how factors were associated with duration for differing subpopulations identified. In addition to traditional risk factors (e.g. herd size, bTB history), we also explored farm geographic area, parcels/farm fragmentation, metrics of intensity via nitrogen loading, and whether herds were designated controlled beef finishing units (CBFU) as potential risk factors for increased duration. The final model fitted log-normal distributions, with two latent classes (k) which partitioned the population into a subpopulation around the central tendency of the distribution, and a second around the tails of the distribution. The latter subpopulation included longer breakdowns of policy interest. Increasing duration was positively associated with recent (<3 yrs) TB history and the number of reactors disclosed, (log) herd size, beef herd-type relative to other herd types, number of land parcels, area, and being designated a controlled finishing unit (“feedlot”), and having high annual inward cattle movements within the “tails” subpopulation. Breakdown length was negatively associated with year of commencement of breakdown (i.e. a decreasing trend) and non-significantly with the organic nitrogen produced on the farm (N kg/hectare), a measure of stocking density. The latter finding may be due to confounding effects with herd size and area. Most variables contributed only moderately to explaining variation in breakdown duration, that is, they had moderate size effects on duration. Herd-size and CBFU had greater effect sizes on the outcome. The findings contribute to evidence-based policy formation in Ireland.

Keywords: Bovine TB; risk factors; disease control; animal health policy; veterinary epidemiology; evidence-based policy

1. Introduction

Bovine tuberculosis remains a priority pathogen of cattle in Ireland (Clarke et al. 2020; Byrne et al. 2020a) and elsewhere (Allen et al. 2018; Gortazar et al. 2011; De Garine-Wichatitsky et al. 2013; De Kantor and Ritacco, 2006), and one that expends considerable resources from farmers, the state, and, for member states, the EU Commission. As an example, direct annual programme costs in Ireland were estimated at €92m in 2018, €47 million from the Exchequer, €9.7 million co-financed by the EU, and €35.2 million paid by farmers (Cawley and Cronin 2019). A key challenge is to ensure that disease policy is evidence-based and adapts in response to changes in risk patterns, so that programme

effectiveness is maximised and stakeholders can be assured that decisions affecting them are underpinned by a coherent epidemiological analysis of the data (More 2019).

The duration of breakdowns can have a significant financial impact on stakeholders (Karolemeas et al. 2010; Milne et al. 2020), and furthermore, prolonged duration breakdowns can be a risk factor for future recrudescence (Byrne et al. 2020b). Therefore, understanding risk factors association with prolonged duration can be of benefit to tailoring policy to local conditions, with the intentions of reducing the spread of infection within and between herds.

In Ireland, a significant expansion of the dairy sector took place from 2015 onwards following the removal of EU milk quotas, with an increase of 54% in milk production from 2005 to 2018 (Kelly et al. 2020); this has been temporally correlated with an increase in TB herd incidence. The increase in dairy production has been associated with an increase in stocking density and herd size, while studies characterised the scale of cattle movements within Ireland (Ashe et al. 2009), it is unclear whether the nature of how these factors relate to TB risk has also changed in that period (Brown et al. 2019). However, data from Britain suggest such movement increases may increase inter-herd *M. bovis* spread (Gilbert et al. 2005).

Duration data are positive integers (natural numbers) data with often unusual distributions which can be difficult to model using conventional statistical means. Previous attempts to model breakdown duration has used categorical models where duration is split into binary outcomes (e.g. Doyle et al. 2016) or ordinal outcomes (e.g. Milne et al. 2020). Alternatively, Milne et al (2020) modelled breakdown duration as a count of days, using a negative binomial random effects model. Time to event modelling (survival) models can be used to model the interval between failures (e.g. Wolfe et al. 2010; Doyle et al. 2016). Alternatively, the duration of time over a fixed period can be modelled by generalised linear models, using various distribution types (e.g. see Moran et al. 2012 for example with hospitalisation length of stay). Breakdown duration, that is the period over which herds are not able to trade due to the presence of TB infection within the herd, in Ireland is both leptokurtic and positively skewed. To explore the variation in breakdown length, a two-class Finite Mixture Model (FMM) was fitted to the data. FMMs are probabilistic models that partition populations into unobserved subpopulations, known as classes or components, without explicitly identifying the subpopulation for which observations belong. Instead, subpopulation membership is modelled as a latent classes, from which the posterior class probability of membership can be estimated, and mixtures of continuous probability distributions of various types (regression models) are used to model the outcome allowing for statistical inference at the subpopulation level. Therefore, FMM are a flexible approach to modelling heterogeneous data (McLachlan and Peel 2004).

The aim of this study was also to investigate different characteristics of herds that may impact on their ability to clear infection, and thus impact breakdown duration so that the controls may be tailored to the risk of persistence of infection. We investigated parameters that have been found to be associated with breakdown duration in other countries, including previous breakdown history, herd type, and herd size, but also explored additional potential predictors. Herd size (i.e. the number of animals on the farm) has been consistently found to be associated with TB breakdown risk, recurrence and duration (Griffin et al. 1996; Olea-Popelka et al. 2004; Green and Cornell 2005; Brooks-Pollock and Keeling 2009; Karolemeas et al. 2010; Skuce et al. 2012; Bessell et al. 2012; Milne et al. 2020). However, herd size could be a proxy for several other farm characteristics (Griffin et al. 1993), for example, the density of animals (e.g. animals per hectare) and the intensity of farming (e.g. high input-high output farms) which could be associated with TB transmission dynamics and maintenance. The geographic area of the farm and configuration (e.g. number of land parcels) could also impact on farm risk of exposure to *M. bovis* from contiguous spread from neighbours or from wildlife sources (Skuce et al. 2010). Therefore, we have added additional proxies for these factors as potential risk factors in our model. This work, therefore, also helps to tease out what aspects of herd size that are associated with breakdown duration in Ireland. Controlled beef finishing units (CBFU), colloquially known as “feedlot” herds, were also identified and modelled as a factor influencing breakdown duration. Previous research elsewhere has shown how such farming systems can be associated with increase prolongation of breakdown duration, associated with increasing genotype

diversity of *M. bovis* facilitated through trade movements (Milne et al. 2019; Milne et al. 2020). Animals from such herds go straight to slaughter, and therefore, are limited in terms of onward interherd transmission but are important to identify in order to explain variation in our outcome of interest. Overall, this study contributes towards informing evidence-based changes to bTB policy to address the risks associated with changing herd dynamics.

2. Results

The final dataset included 17,312 breakdowns. The mean breakdown length was 177.85 days (SD: 105.25; Median: 147; IQR: 135-187; Fig. 1). The distribution of breakdown durations was both leptokurtic (kurtosis: 68.24) and positively skewed (skewness: 5.90; Fig. 1).

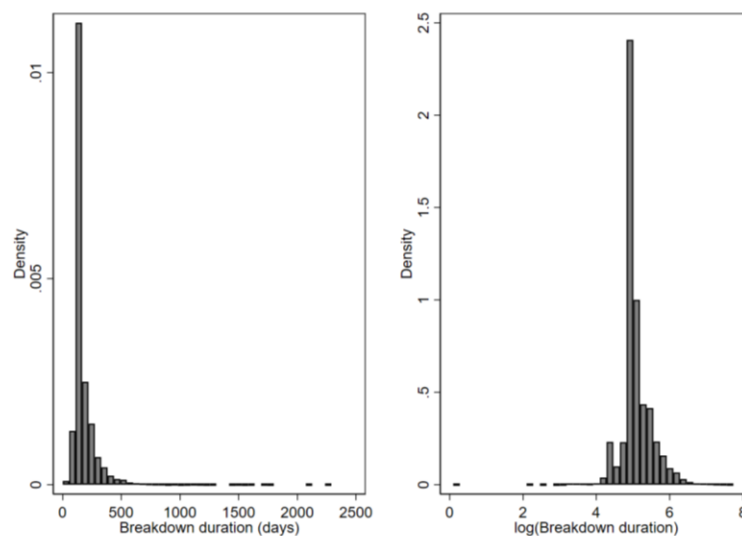


Figure 1. Breakdown duration days (left) and with duration log-transformed (right).

2.1 Model distribution and class probability

Distribution fits to null models (that is, without covariates) were compared using BIC and AIC, with the following FMM models compared (Fig. 2): 1. Ordinary Least Squares (OLS) log-linear; 2. GLM lognormal; 3. Gamma distribution; 4. Weibull distribution (Table 1). GLM Poisson and Negative binomial models did not converge. OLS models, where the outcome was natural logarithm transformed, were best fitting models with smallest AIC and BIC. There was evidence also that FMM models (with $k > 1$) were favoured in comparison with a model without latent subpopulation classes.

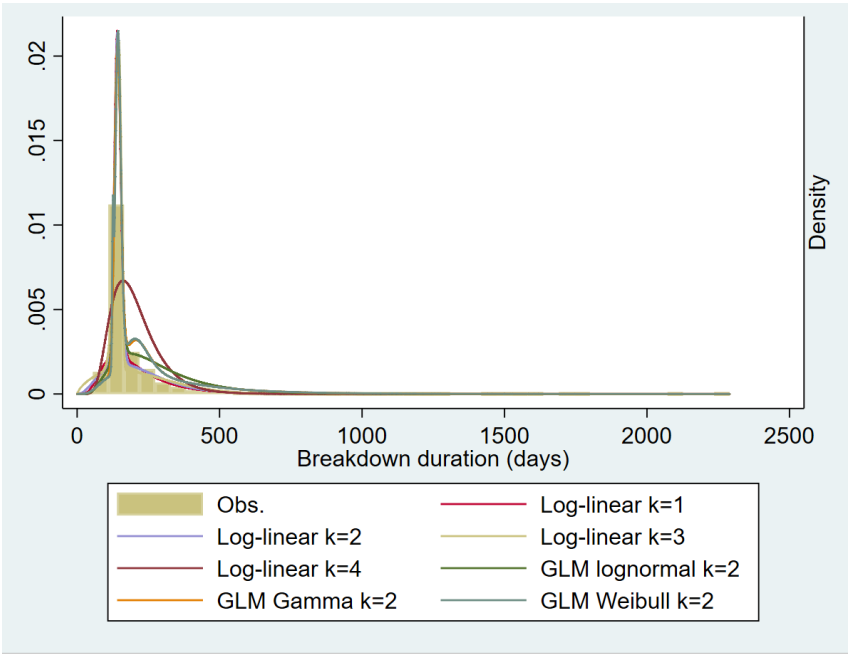


Figure 2. A comparison of FMMs with differing distributions to model classes (OLS log-linear k=1; OLS log-linear k=2; OLS log-linear k=3; GLM lognormal k=2; GLM gamma distribution model k=2; GLM Weibull distribution model k=2)

Comparing the number of classes (k) to fit to the data, a three-class and four-class model fitted the data better, with a lower BIC and AIC relative to the two-class model (Table 1). However, the parsimony of the two class models were preferred. Furthermore, research has shown the overfitting may be an issue for mixture models where $k > 2$, and convergence for more complicated models can be an issue for mixture models (Magussen et al. 2019). The best fitting model with $k=2$ was the log-linear regression model [outcome variable was log-transformed before fit using the REGRESS command].

Table 1. Comparison of Bayesian and Akaike’s Information Criterion values for finite mixture models fitted to bovine TB breakdown duration data. k = number of latent classes (subpopulations)

Model	BIC	AIC
Log(duration), k=1	17678.83	17663.32
Log(duration), k=2	6532.47	6493.67
Log(duration), k=3	6092.97	6030.90
Log(duration), k=4	5737.85	5652.50
GLM lognormal, k=2	182664.06	182625.26
GLM gamma, k=2	183161.31	183122.52
GLM Weibull, k=2	185397.98	185359.19

The two latent distributions fitted to different parts of data distribution (Fig. 3). Figure 3 presents the predicted posterior probability of class membership for class 1 (black line) and class 2 (blue). The probability of class 1 membership (relative to class 2 membership) was very high for very large and

small values. Class 2 membership was limited primarily to a narrow range around the central tendency. These data are visualised in figure 4 by plotting the histograms of data with >20% probability of membership for PR2 class (clear) and PR1 class (red) (Fig. 4).

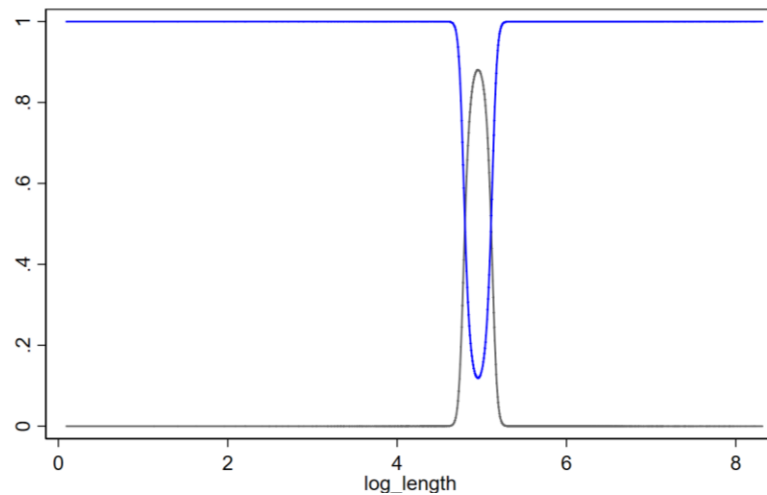


Figure 3. The predicted probability of class membership for a two-latent class FMM null model fitted to bovine TB breakdown duration data.

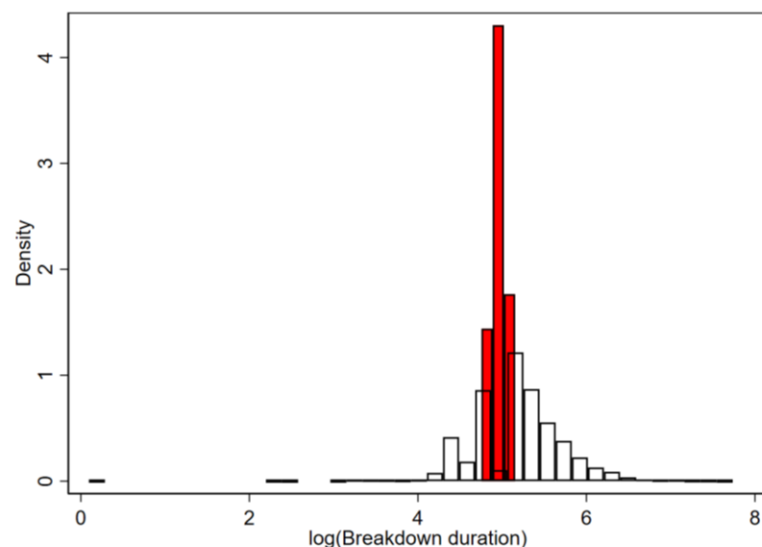


Figure 4. The predicted density from an FMM with two classes fitted to bovine TB breakdown duration data; class 1 is represented as red bars which cluster around the central tendency of the distribution; class 2 is represented as clear bars in the histogram which encapsulated more of the variance in the dataset away from the mean.

2.1 Development of a multivariable log-normal FMM

A baseline model was developed with fixed set of covariates log-herd size, herd-type, and herd history (breakdown history over previous 3 years), as potential predictors of log-breakdown length based on previous research. The probability of being class/component 1 or 2 was modelled as a function of log-herd size, herd history and herd type, as we hypothesized that breakdown duration may be associated with these factors. Comparisons were made using BIC and AIC with baseline models with the inclusion or exclusion of these factors, separately and respectively. Baseline models

had the lowest BIC values when all three factors were retained to explain variation in breakdown length, and when both log-herdsize and herd type were retained as factors explaining some variation in the probability of being associated with class 1 or 2 (Baseline vs. class prob(log_hs, herd_type) ΔBIC: 43.75; Table 2). The model with the lowest AIC included herd-history as an additional predictor of class probability, but the more parsimonious (BIC) model was selected.

Table 2. Comparison of base FMM models using Bayesian and Akaike’s Information Criteria.

Model	Class probability	Regression Predictors	BIC	AIC
Log(duration), k=2	.	log_hs herd_history herd_type	5893.03	5730.09
Log(duration), k=2	log_hs	log_hs herd_history herd_type	5858.36	5687.66
Log(duration), k=2	herd_type	log_hs herd_history herd_type	5853.46	5667.24
Log(duration), k=2	herd_history	log_hs herd_history herd_type	5924.37	5730.39
Log(duration), k=2	log_hs, herd_type	log_hs herd_history herd_type	5849.29	5655.31
Log(duration), k=2	log_hs, herd_history	log_hs herd_history herd_type	5884.57	5682.83
Log(duration), k=2	herd_type, herd_history	log_hs herd_history herd_type	5879.33	5662.08
Log(duration), k=2	log_hs, herd_type, herd_history	log_hs herd_history herd_type	5872.54	5647.52

Additional predictors of breakdown duration were added to this model. The model fitted a controlled beef finishing unit proxy “feedlot” status (binary; 1 if herd ever had designation as a controlled beef finishing unit; 0 otherwise), an intensity metric based on nitrogen loading per hectare (logarithm of median bovine nitrogen per hectare over 2015-2018), the number of parcels (recorded per herd in 2018 using LPIS at a resolution of 7m), land area (median hectares over 2015-2018), and breakdown start year, and inward animal movement metric. The mean number of inward animal moves (2013-2018) into a farm was dichotomised into ‘large’ (mean >=90 per annum) and other (<90 mean moves).

The final model outputs are presented in tables 3 & 4. Predictions from the model (Figure S1) showed that class 1 were tightly bound around the central tendency of the distribution, “central” subpopulation; class 2 had greater representation at the tails of the distribution, “tail” subpopulation. The entropy value, which discerns how distinct subpopulation classes were, was estimated at 0.58, which is moderate but reasonable for a large population size with 2-classes (Jaki et al. 2019). The subpopulations were reasonably balanced, with 51% of the population classified as Class 1 “central” subpopulation, and 49% being falling into the Class 2 “tail” subpopulation. The (geometric) mean breakdown duration for Class 1 subpopulation was 142 days (exp(4.96)), while subpopulation Class 2 the mean duration was 183 days (exp(5.21)).

Table 3. Concomitant variables as predictors of class/component probabilities; coefficients represent associations with Class 2 relative to Class 1.

	Coef.	exp(Coef.)	SE	z	P	Lower 95%	Upper 95%
log_hs	0.070	1.073	0.018	3.860	<0.001	0.035	0.106
herd_type_num							
Beef	<0.001	1.000					
Dairy	0.158	1.171	0.061	2.590	0.010	0.039	0.277
Other	-0.331	0.718	0.099	-3.350	0.001	-0.525	-0.137
Suckler	-0.096	0.909	0.051	-1.870	0.061	-0.196	0.005
Contant	-0.317	0.728	0.080	-3.970	<0.001	-0.473	-0.161

The preferred model with lowest BIC retained both log-herd size and herd type as explaining variation in the probability of herds being associated with either latent class component. Table 3 shows the output of this part of the mixture model. The model suggested that increasing log-herd size was significantly associated with having greater probability of being in Class 2 “tail” subpopulation. Dairy herds were significantly greater probability of being represented in Class 2 “tail” component, relative to Class 1 “central”, in comparison with beef rearing/finishing production type ($p=0.010$). ‘Other’ herd types had lower probability of being in class 2 relative to beef herd types ($p<0.001$). There was no significant difference in class probabilities between beef rearing/finishing and beef suckler ($p=0.876$) herd type.

The associations between factors hypothesized to impact breakdown duration for Class 1 and Class 2 are presented in Table 4. All predictors were retained in the log-normal regression parts of the model explaining breakdown duration for both component/class 1 and 2. Comparing coefficients between Class 1 and Class 2 suggests that risk factors may have different associated effect on breakdown duration.

Class 1 breakdowns duration tend to decrease with increasing nitrogen load per hectare (log; **log_NPH**) and for breakdowns commencing later in the time series (**bd_start_yr**). Duration was positively associated with herd history (**herd_history**), with herds with breakdowns during the previous three years having a longer duration than those without breakdowns. There was not, however, a clear dose-dependency in terms of breakdown size (i.e. number of reactors disclosed). Herds with 1, 2-4, or >4 reactors were significantly higher risk for longer breakdowns within this ‘central’ subpopulation, relative to herds without reactors. There was no difference in breakdown duration between herds without a previous breakdown and ‘zero-reactor’ breakdowns (i.e. historic breakdowns that were disclosed due to a TB lesion detected at routine slaughter and in which no additional reactors were disclosed at follow-up herd tests). There was no significant association between duration and herd size, CBFU (feedlot) status or farm fragmentation (**parcel_cut**). Inward movements were dichotomised into herds which had high mean yearly inward movements (>90 moves) or lower mean inward movements (<90). In Class 1 subpopulation, there was a negative association between high inward movements and breakdown length (-0.039; 95%ci: -0.04 - -0.032; 3.8% reduction in duration).

Table 4. Estimated associations between factors and breakdown duration (log_length) for Class 1 and Class 2 subpopulations from a finite mixture model.

log_length	Class 1 "Central"				Class 2 "Tail"			
	Coef.	P	Lower 95%	Upper 95%	Coef.	P	Lower 95%	Upper 95%
log_hs	<0.001	0.835	-0.003	0.002	0.090	<0.001	0.077	0.102
herd_history								
TB free	ref				ref			
Zero reactors	-0.001	0.787	-0.009	0.007	0.115	<0.001	0.068	0.162
1 reactors	0.015	<0.001	0.007	0.023	0.129	<0.001	0.076	0.182
2-4 reactors	0.017	<0.001	0.009	0.025	0.219	<0.001	0.166	0.273
>=5 reactors	0.017	0.004	0.005	0.029	0.226	<0.001	0.155	0.296
herd_type_num								
Beef	ref				ref			
Dairy	0.025	<0.001	0.018	0.031	-0.075	<0.001	-0.114	-0.036
Other	0.003	0.560	-0.006	0.012	-0.095	0.003	-0.157	-0.033
Suckler	0.006	0.015	0.001	0.011	-0.014	0.373	-0.046	0.017
Controlled finishing units								
Never	ref				ref			
Designated	-0.001	0.932	-0.030	0.027	0.710	<0.001	0.603	0.818
log_NPH	-0.003	0.040	-0.005	<0.001	-0.014	0.054	-0.028	<0.001
parcel_cut								
1	ref				ref			
2	<0.001	0.902	-0.006	0.006	0.026	0.177	-0.012	0.063
3	0.004	0.224	-0.002	0.010	0.031	0.110	-0.007	0.069
4	0.005	0.116	-0.001	0.012	0.064	0.002	0.023	0.105
>=5	0.001	0.736	-0.005	0.007	0.049	0.006	0.014	0.084
med_area	<0.001	<0.001	<0.001	<0.001	<0.001	0.299	<0.001	0.001
cut_mmoves								
<90	ref				ref			
>90	-0.039	<0.001	-0.047	-0.032	0.108	<0.001	0.062	0.155
bd_start_yr	<0.001	0.470	-0.001	0.001	-0.023	<0.001	-0.029	-0.017
Contant	4.234	<0.001	2.247	6.221	51.278	<0.001	39.442	63.114

Herds associated with class 2 'tail' subpopulation exhibited significant increased breakdown duration with increasing herd-size (**log_hs**), and if designated with a 'feedlot' status (CBFU). Duration of breakdowns increased by 9.4% (β : 0.090; 95%CI: 0.077-0.102) for every log-unit increase in herd size. Feedlot herds within class 2 subpopulation were 203.4% (β : 0.710; 95%CI: 0.603-0.818) longer duration relative to non-feedlot herds. The median predicted duration for feedlot herds was 451 days, whereas as the median duration was 182 days for non-feedlot designated herds. The predicted relationship for this subpopulation for herd size and feedlot designation is presented in Fig.5.

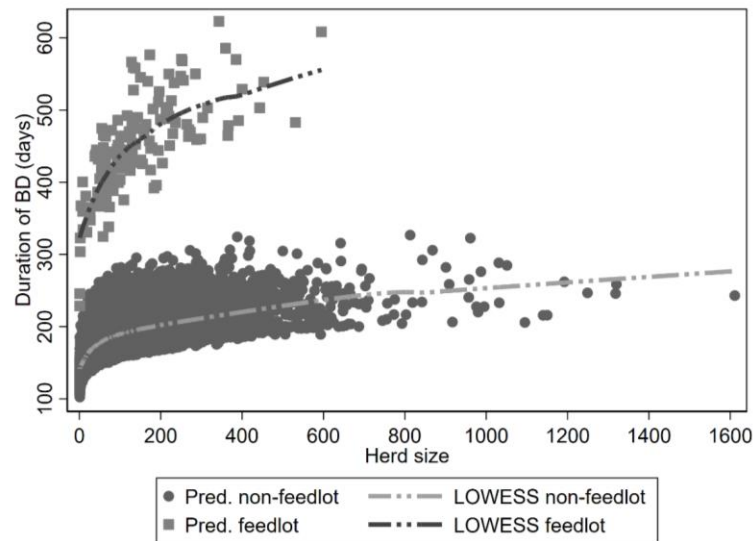


Figure 5. Relationship between herd size and breakdown duration as predicted for Class 2 subpopulation of a finite mixture model. Squares are predictions for herds with a controlled beef finishing unit “feedlot” designation; Points are predictions for herds without “feedlot” designation. Lines are Local Weighted Scatterplot Smoothing (LOWESS) curves fitted to the data.

Duration increased if a herd had previously experienced a breakdown during the previous 3-years. There was a trend towards dose-dependency, with increasing duration associated with greater number of reactors disclosed during the previous breakdown (Fig. 6), however the size effect was reasonably small. For example, there was a 25% (β : 0.226; $p < 0.001$; 95%CI: 0.155-0.296) increase in mean duration for herd with five or more reactors relative to herds which did not experience a breakdown during the previous 3 years.

Parcel number (parcel_cut), a measure of farm fragmentation, was positively associated with increased breakdown length, though only for herds with 4 or more parcels relative to herds with 1. Again, effect size was moderate, for example, a 6.6% (β : 0.064; $p < 0.001$; 95%CI: 0.023-0.105) increasing duration for herds with 4 parcels relative to herds with only 1 parcel.

Beef rearing/finishing herds had significantly shorter breakdown duration relative to dairy, suckler or “other” herd types (herd_type) (Table 4; Fig. 6). For example, dairy herds experienced mean breakdown log-lengths that were 7.2% shorter ($\exp(\beta) - 1$) relative to beef herds.

The linear predictor for start year of breakdown (bd_start_yr) suggested that breakdown lengths decreased slightly per year of the study for Class 2 subpopulation (0.5% reduction per year; β : -0.023; $p < 0.001$; 95%CI: -0.029 - -0.017).

Controlling for herd size, intensity metric based on nitrogen loading (NPH) was negatively associated with breakdown length, though this effect was not significant (β : -0.014; $p = 0.054$; 95%CI: -0.028 - <0.001). It was likely that the effect was impacted by confounding with herd size and area.

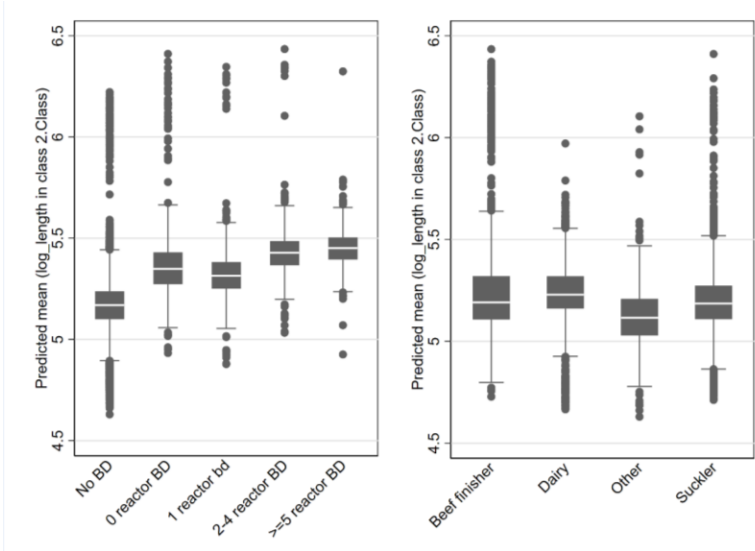


Figure 6. Relationship between breakdown history during previous 3 years and predicted mean breakdown duration (top). Relationship between herd type and predicted mean breakdown duration (bottom). Note, log scale.

In Class 2 subpopulation, there was a positive association between breakdown length and high inward movement (0.108; 0.062-0.155; 11.4% increase in duration). The size effects for both subpopulations are reasonably small, for example, herds in subpopulation 2 increase their mean duration by approximately 20.8 days if they make >90 inward moves per annum (i.e. mean[183 days]*0.114).

3. Discussion

This study was the first to explore variation in breakdown duration explicitly in Ireland, and more generally the first to use finite mixture models (FMM) to explore the variation in breakdown duration for bTB. It is also the first study to investigate cattle production of nitrogen per hectare (a common measure of the intensity of livestock production within the EU) as a potential factor associated with bovine TB breakdown duration. The significance of this approach lies in enabling the adaption of disease policies to changing risks based on an evidence base which is inclusive of larger trends such as livestock intensification.

Previous research has dichotomised breakdown duration and modelled using logit distributions (e.g. Reilly and Courtenay 2007; Doyle et al. 2016; Karolemeas et al. 2010; Milne et al. 2019a), have broken down duration into ordered categories of increasing duration and modelled using ordinal regression models (Milne et al. 2020) or have modelled duration as a count of days using a negative binomial discrete distribution (Milne et al. 2020) or using simulation approaches (Brooks-Pollock et al. 2009). The durations recorded in the present study exhibited an unusual distribution, with both positive skewness and kurtosis. The mixture model allowed for two different distributions to be fit to the data, with class associations being modelled via a latent class model.

The latent classification of the data into two classes suggested that there were observations clustering tightly around the central tendency, and a second which captured more of the variation in the outcome, especially at the tails of the distribution. The latent component of the model suggested that larger herds and dairy, relative to beef, herds had a higher probability to be represented by the tail subpopulation, relative to the central subpopulation. This may indicate that such herd characteristics are associated with more unusual breakdowns away from the central tendency. Previous research has found that chronic breakdowns, which would be represented by our tails subpopulation, can be associated with both increasing numbers of animals within the herd (e.g. Karolemeas et al. 2010) and dairy enterprises (Reilly and Courtenay 2007; Milne et al. 2018) relative to other herd types.

In terms of interpretation, the latent classification and the “tail” subpopulation log-normal regression model are informative. The log-normal regression for the ‘central’ subpopulation is less informative, from a policy perspective, as: 1. It seeks to explain variation over a narrow range; that is, the variance in the outcome was low, 2. Larger breakdowns are of greater significance from a disease policy development perspective, assuming that at least partially the speed at which average breakdowns clear may be more related to breakdown management/compliance and response times (e.g. actions by the farmer and operational delivery by veterinarians) as well as epidemiologically relevant factors. However, it is interesting to note the differences in the relationship between outcome and predictors between subpopulations. In the central subpopulation there was no significant association with duration and herd-size, number of land parcels, controlled finishing units (feedlots) or breakdown starting year. Dairy herds varied in their association with breakdown duration across subpopulations, with longer durations relative to beef herds in subpopulation 1, and shorter durations relative to beef herds in subpopulation 2. However, it should be borne in mind that dairy herds have a higher class-probability for subpopulation two. Furthermore, dairy herds tend to be larger and have older age profile animals, both of which are associated with bTB risk (Brooks-Pollock and Keeling, 2009; Brooks-Pollock et al. 2013).

The class 2 log-normal part of the FMM suggested that breakdown lengths were significant increased with increasing numbers of reactors disclosed at breakdown during the previous 3 years. Herd history has been associated with breakdown length in previous studies outside of Ireland, though metrics of TB history were not always analogous to what was done in the present study (Doyle et al. 2016; Milne et al. 2020). A history of TB in a herd was associated with large (>13 reactors) bTB breakdowns, relative to small (2-4 reactors) in a case-control study in Ireland (Clegg et al. 2018). Interestingly, recent work suggested that that previous breakdown size was not associated with recurrence risk in Ireland (Byrne et al. 2020b).

The factors that appeared to have the largest effects on breakdown duration were herd-size and whether the herd was ever designated a controlled beef finishing unit ‘feedlot’ herd. Controlled

finishing unit herds are non-breeding beef finishing units which are only permitted to move cattle directly to slaughter and fulfil certain criteria to receive the designation, for example, biosecurity protocols are in place, and boundaries are double fenced or equivalent to prevent any direct contact with cattle on neighbouring lands/premises/holdings. Such herds may apply for controlled finishing unit herd designation upon breakdown, subject to there being no evidence of within-herd spread; breakdowns are typically due to lesions detected at routine slaughter from cattle bought-in and fattened over a relatively short period of time for finishing. This status enables them to continue to buy in beef cattle for finishing during breakdowns. These herds are deemed of lower risk for onward infection transmission due to high turnover of animals within the herds before going direct to slaughter. The effectiveness of controls to prevent such herd enterprises posing a risk to neighbouring herds via wildlife spillover or contiguous spread is subject to another ongoing study (J. Madden pers. comm.). Beef fattening herds in Northern Ireland, which can buy-in animals during breakdowns, have increased strain diversity of *M. bovis* and increased breakdown duration time, linked to inward animal movements from a wide geographical area (Milne et al. 2019b; Milne et al. 2020). The association between such herds and increased breakdown length identified in this study and in Northern Ireland may be explained by the ability of the business model to continue while under restriction; the length of breakdown does not reflect a chronic disease problem on these farms *per se* but rather the reality that a farm practice which involves purchasing in large numbers of beef cattle for fattening and slaughter makes it likely that the detection of bTB lesions at routine slaughter would be a regular occurrence, given the background level of bTB in Ireland. Each such detection extends the breakdown again, lengthening breakdown duration but not necessarily reflecting the continuing spread of infection within the herd.

Herd size has been found to be a significant factor for TB breakdown risk in general for many studies (e.g. see Skuce et al. 2010; Allen et al. 2018). Brooks-Pollock and Keeling (2009) showed how herd size was associated with breakdown persistence, which the authors interpreted as indicating a lack of clearance of infection within herds after a breakdown. We have attempted to explore what correlates and related variables of herd size that might be important for breakdown duration with the inclusion of nitrogen load (a measure of agricultural intensity, based on organic nitrogen per hectare), land area, and parcel number. For Class 2, “tail” subpopulation, there were significant increases in log-length of breakdown with increasing land area and the number of parcels (primarily 4 or more parcels, relative to single parcel units), while controlling for herd size. However, for both factors, the effects size was small. Despite this, previous work elsewhere has found significant positive associations between farm fragmentation (Johnston et al. 2005; Mill et al. 2012; Broughan et al. 2016; but see Griffin et al. 1993; Vial et al. 2011) and land area (Mill et al. 2012; Vial et al. 2011) and TB breakdown risk, though there are little comparative data on these factors relationship with breakdown duration.

Surprisingly, controlling for herd size there was a negative relationship with increasing nitrogen load. This outcome is mostly likely affected by significant confounding relationship between herd size, area and nitrogen load (see Figure S2). Indeed, nitrogen per hectare (NPH) was positively and significantly associated with breakdown duration at the univariable level and for a model excluding herd size and area confounders (Figure S3; Table S1). The final multivariable model suggested that herd size, that is the number of animals/potential hosts, appears to be the key herd-level factor for predicting breakdown length.

It was encouraging that for both subpopulations that the temporal trend was for a reduction in mean duration over time. This suggests that interventions aimed at clearing infection once reactors are identified are being more effective over time. However, it should be noted that the estimated reduction is very small, being less than a day reduction in mean duration for the class 2 subpopulation. Previous work in Ireland has found a decreasing trend also in breakdown recurrence from 1998-2015 (Houstma et al. 2016), though recent trends suggests incidence is increasing.

In conclusion, our work has suggested breakdown duration in Ireland is related to several herd and farm characteristics, including the herd size, type, area (acreage), fragmentation, and trading

frequency. Given that herd size and acreage contributed additively to increasing breakdown duration, trends in farming including farm expansion and consolidation may signal future increases in breakdown duration, causing greater challenge to the bTB control program.

4. Materials and Methods

4.1 Data

The dataset used is a breakdown herd dataset where each line in the dataset represents a discrete breakdown (Byrne et al. 2020). Throughout, the start date and end date within that file was used to define breakdown durations (see Byrne et al. 2020b for additional detail). All breakdowns were initiated by the disclosure of (i) bTB reactors from Single Intradermal Comparative Cervical Tuberculin test (SICCTT) or (ii) the disclosure of a bTB lesion found at slaughter (factory lesion). End dates were when herds were derestricted, in compliance with the EU regulations on bovine TB. The dataset included breakdowns that were ongoing between Jan 2012 and Dec 2018 and were initiated from Jan 2012-Dec 2018. Follow-up period ended in August 2019.

Two inclusion criteria were employed to the data. Only the most recent breakdowns per herd were selected; this was to avoid repeated measures per herd, and therefore simplify the analysis. All breakdowns that were censored by follow-up (i.e. they were still experiencing a breakdown when follow-up finished in August 2019) were removed from the dataset.

4.2 Model

Finite Mixture Models (FMM) are a flexible modelling framework that have been applied to several biological problems, including clustering, estimation of diagnostics performance, and for modelling heterogeneous data (McLachlan and Peel 2004). In addition, FMM is an approach that can be applied to duration models (McLachlan and McGiffin 1994; McLachlan and Peel 2004; Moran et al. 2012). Applications of FMM model to duration include modelling length of stay in hospitals (Moran et al. 2012) and age-at-onset for various conditions in human medicine including migraine or bipolar disorder (e.g. Asuni et al. 2010; Manchia et al. 2010; Nowrouzi et al. 2015). FMM gives flexibility to fit to the data by allowing admixture of distributions, called classes (k) to explain the variation in the data. Furthermore, different continuous probability distribution types can be fitted, including normal, log-normal, Poisson, negative binomial, Gamma and Weibull. For example, log-normal distributions have been applied to 'survival' type data, as an alternative to Cox proportional hazard models (Royston 2001). Here we developed a model of breakdown duration without repeated measures using an FMM approach.

Models were compared using Bayesian information criteria (BIC) and Akaike's Information Criteria (AIC) values, which favour models which explain variation in our outcome, while penalising models with many vaguely informative parameters. Where BIC and AIC did not support the same model configuration, the model with least parameters were preferred, which was generally the BIC preferred model. Simulation studies has shown that BIC is generally preferred to AIC for FMM, as AIC can tend to select for too many classes (Van Horn et al. 2012).

Entropy value, which discerns how distinct subpopulation [latent] classes were, was estimated (Jaki et al. 2019). The higher the number the greater the separations between classes, 'rule of thumb' threshold of 0.8 has been proposed. However, recent simulation analysis has shown that entropy values are sensitive to sample size. Entropy values <0.2 might be expected in large sample studies (e.g. >3000 observations, 1500 per subpopulation) with balanced split between 2-classes and where true sample separation has been simulated (Jaki et al. 2019).

Our process was to plot out the distribution of breakdown durations and fit differing FMM types, including log-normal, normal, and count distributions. We explored $k>2$ class models, but only

retained a two-class model for multivariable models to ensure a more parsimonious model that was more meaningful from an epidemiological perspective. Furthermore, higher number of components in FMMs are associated with overfitting (Magnussen et al. 2019).

The steps involved in the construction of the model is depicted in figure S4. Once we developed a well-fitting null model without covariates, we built multivariable FMMs. Initially, baseline models were fit with the parameters: herd size, herd history, and herd types. Then additional parameters were added sequentially. Comparisons between model configurations were made using AIC/BIC values; lower values indicated better performing models. Outcome and covariate parameter summary statistics are presented in table 5.

Table 5. Summary statistics of outcome and covariate parameters for FMM model of bTB breakdown duration

Variable	n	Mean	Std. Dev.	exp(mean)	Min	Max
log_length	17312	5.087	0.403	161.869	0.095	7.736
log_hs	17312	4.058	1.162	57.887	<0.001	7.385
log_NPH	17312	4.397	0.986	81.210	-4.605	6.733
med_area	17312	49.360	39.962		<1.000	754.900
herd_history						
TB free	13833	80%			0	1
Zero reactors	1156	7%			0	1
1 reactor	896	5%			0	1
2-4 reactors	935	5%			0	1
>=5 reactors	492	3%			0	1
herd_type						
Beef	3576	21%			0	1
Dairy	4743	27%			0	1
Other	830	5%				
Suckler	8163	47%			0	1
Feedlot						
0	17173	99%				
1	139	1%			0	1
parcel_cut						
1	3030	18%				
2	3299	19%			0	1
3	3069	18%			0	1
4	2425	14%			0	1
>=5	5489	32%			0	1
cut_mmoves						
0	15660	90%				
1	1652	10%			0	1
bd_start_yr						
2012	1231	7%				
2013	2473	14%			0	1
2014	2466	14%			0	1
2015	2486	14%			0	1
2016	2571	15%			0	1
2017	2935	17%			0	1
2018	3150	18%			0	1

Herd size was log-transformed. Herd history was a categorical variable indicating whether the herd had previously had a breakdown within 3-years, and if so, the size of the breakdown in terms of reactors disclosed (0, 1, 2-4, ≥ 5). Breakdowns with 0 reactors would have been initiated by lesions found during slaughter at abattoir during post-mortem surveillance (Byrne et al. 2020). Herd types were taken from the AHCS dataset, as beef, dairy, 'other' or suckler herds. Herd type was based on standard definitions used by DAFM: dairy herds are self-explanatory; beef herds are non-breeding herds; suckler herds are beef breeding herds; the designation "other" reflects herds which engage in a combination of such activities.

Other variables offered to the models included: controlled finishing unit (colloquially referred to in Ireland as "feedlots") status, nitrogen loading, number of farm fragments/parcels, number of moves into the herd, and year of breakdown initiation.

"Feedlot" status was assigned if a herd was ever designated a controlled beef finishing unit (CBFU) by DAFM at breakdown from 2006-2018 (see Discussion).

Nitrogen, measured as the organic nitrogen loading per hectare (NPH) from bovines, was measured as the median NPH per herd over 2015-2018 assigned to each herd. Because of the skew of this variable (small number of herds with very large NPH values) the parameter was log-transformed. These data were estimated by DAFM in accordance with Nitrogen Regulations in Ireland.

The number of farm parcels was measured from the Land Parcel Identification System (LPIS) 2018, with discrete land-parcels being classified if parcels were $>7\text{m}$ from each other. This buffer was chosen to ensure that parcels that effectively were single units were not counted as separate parcels. Land parcels were categorised into 1, 2, 3, 4 or ≥ 5 parcels.

Land area was measured as the mean hectares land area declared during 2015-2018 for Basic Payment Scheme (BPS) purposes. The median of these values per herd was assigned as a metric of geographic footprint for each herd in the dataset. There was a 99% correlation between individual yearly and median land area.

Movement of animals into herds were measured by a single metric in this instance, as the mean of total inward moves per annum from 2013-2018, inclusive. Univariable models suggested that most variation in the relationship between breakdown length and inward moves was whether herds brought in large numbers of cattle or not. Therefore, the data were categorised into herds with an average of >90 inward moves per year and those with <90 inward moves per year.

Year of breakdown initiation was entered into models initially as categorical variables but added many parameters ($n=7$ parameters; 2012-2018). Exploratory analysis suggested that trends could be modelled as a single linear predictor.

All analysis was undertaken in Stata 15.1 MP – Parallel Edition (Stata Corp., 2017) using the FMM suite of commands.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Output plots of the class probability and predicted log(breakdown duration) from a FMM of bTB breakdowns from a full model, Figure S2: Relationship between harm area (Ha) or herd size (number of cattle) and total bovine organic nitrogen per farm (kg). Figure S3: Relationship between predicted duration and log(NPH) for non-feedlot herds from a FMM model excluding log-herd size and farm area (ha). Figure S4: Schematic of model construction steps. Table S1: FMM log-normal parameter estimates for a model excluding log-herd size and farm area (ha) with two-components/classes. Class probabilities were modelled with variation across herd types.

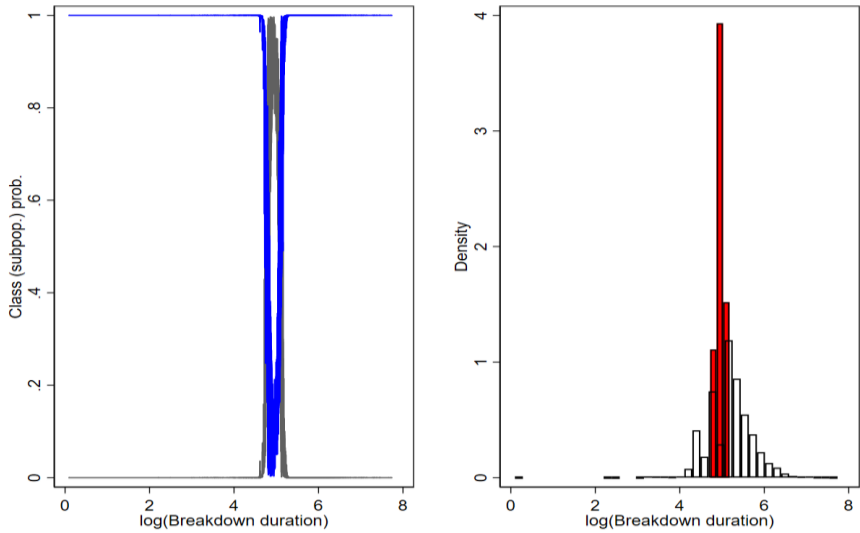


Figure S1: Output plots of the class probability (left panel) and predicted log(breakdown duration) from a FMM of bTB breakdowns from a full model (right panel).

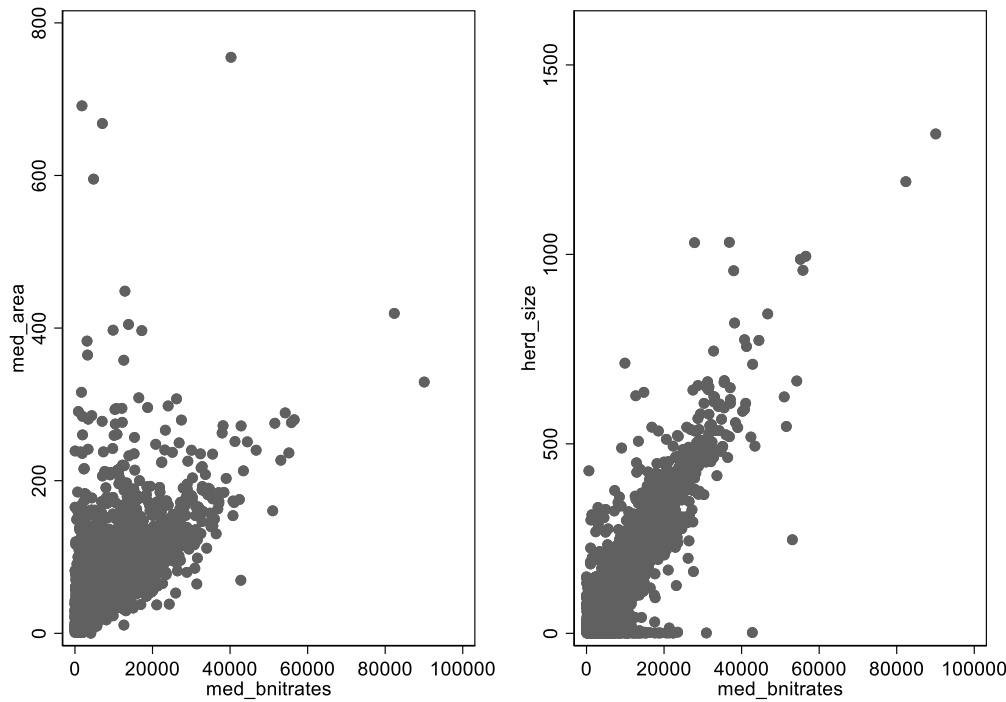


Figure S2: Relationship between harm area (Ha) or herd size (number of cattle) and median bovine organic nitrogen per farm (kg).

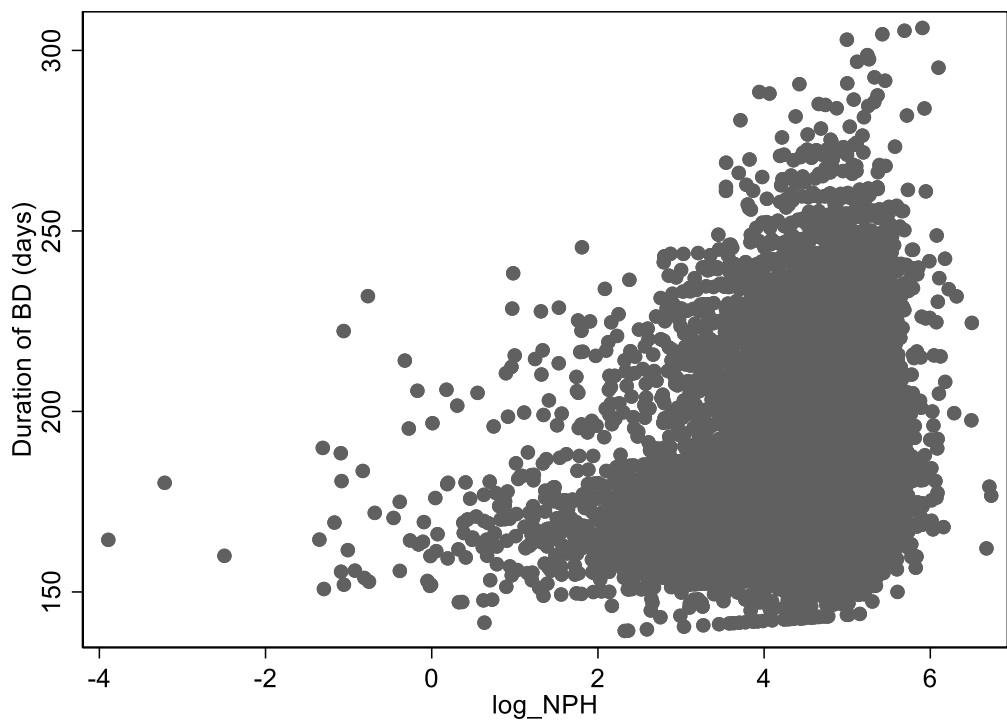


Figure S3: Relationship between predicted duration and log(NPH) for non-feedlot herds from a FMM model excluding log-herd size and farm area (ha)

Table S1: FMM log-normal parameter estimates for a model excluding log-herd size and farm area (ha) with two o-components/classes. Class probabilities were modelled with variation across herd types.

Outcome: log length of breakdown								
	Class 1 "Central"				Class 2 "Tail"			
	Beta	p	Lower 95%	Upper 95%	Beta	p	Lower 95%	Upper 95%
herd_history								
TB free	ref							
Zero reactors	-0.002	0.624	-0.010	0.006	0.135	0.000	0.088	0.183
1 reactors	0.014	0.001	0.006	0.022	0.151	0.000	0.098	0.205
2-4 reactors	0.016	0.000	0.008	0.024	0.250	0.000	0.196	0.304
>=5 reactors	0.014	0.015	0.003	0.026	0.267	0.000	0.195	0.339
herd_type_num								
Beef	ref							
Dairy	0.019	0.000	0.013	0.026	0.020	0.306	-0.018	0.057
Other	0.002	0.722	-0.007	0.011	-0.077	0.015	-0.139	-0.015
Suckler	0.005	0.040	0.000	0.010	0.002	0.909	-0.030	0.034
Feedlot								
Never	ref							
Designated	-0.001	0.939	-0.030	0.028	0.737	0.000	0.628	0.846
log_NPH	-0.002	0.092	-0.004	0.000	0.012	0.076	-0.001	0.025
parcel_cut								
1	ref							
2	-0.001	0.855	-0.007	0.006	0.036	0.060	-0.002	0.074
3	0.003	0.313	-0.003	0.010	0.051	0.009	0.012	0.089
4	0.004	0.210	-0.002	0.011	0.091	0.000	0.050	0.132
>=5	-0.002	0.577	-0.007	0.004	0.095	0.000	0.061	0.129
cut_mmoves								
<90								
>90	-0.045	0.000	-0.052	-0.038	0.178	0.000	0.133	0.223
bd_start_yr	0.000	0.637	-0.001	0.001	-0.022	0.000	-0.028	-0.016
_cons	4.480	0.000	2.484	6.476	49.086	0.000	37.088	61.083

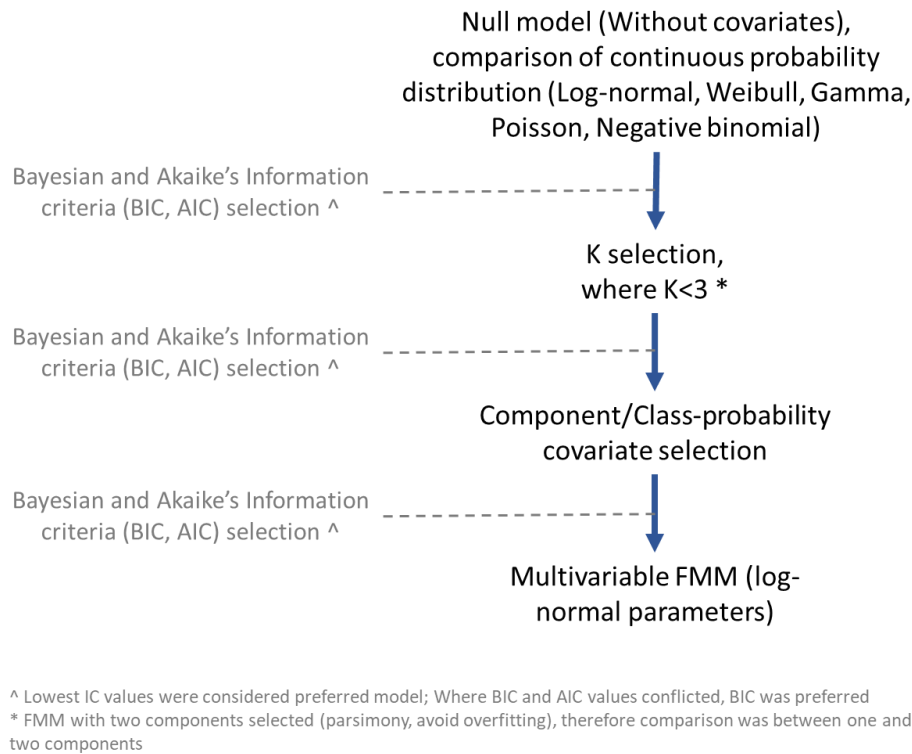


Figure S4: Schematic of model construction steps.

Author Contributions: Conceptualization, AWB, PB, ER; methodology, AWB; formal analysis, AWB, JM; investigation, AWB, JM; data curation, JM; writing—original draft preparation, AWB; writing—review and editing, AWB, DB, JOK, JM, PB, ER; supervision, DB. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to acknowledge the support of the One-health Scientific Support Unit and the provision of nitrogen data by colleagues in the Nitrates, Biodiversity and Engineering Division, DAFM.

Conflicts of Interest: The authors declare no conflict of interest. While there was no external funding secured for this research, the senior authors of this paper are employees of the Department of Agriculture, Food and the Marine, Government of Ireland, however the decision to publish rested with the authors.

References

- Allen, A.R., Skuce, R.A. and Byrne, A.W., 2018. Bovine tuberculosis in Britain and Ireland—A perfect storm? the confluence of potential ecological and epidemiological impediments to controlling a chronic infectious disease. *Frontiers in Veterinary Science*, 5, p.109.
- Ashe, S., More, S.J., O'Keeffe, J., White, P., McGrath, G. and Aznar, I., 2009. Survival and dispersal of a defined cohort of Irish cattle. *Irish Veterinary Journal*, 62(1), pp.1-6.
- Asuni, C., Manchia, M., Deidda, A., Stochino, M.E., Cherchi, A. and Del Zompo, M., 2010. Mixture analysis of age at onset in migraine without aura: evidence for three subgroups. *Headache: The Journal of Head and Face Pain*, 50(8), pp.1313-1319.
- Bessell, P.R., Orton, R., White, P.C., Hutchings, M.R. and Kao, R.R., 2012. Risk factors for bovine tuberculosis at the national level in Great Britain. *BMC Veterinary Research*, 8(1), p.51.
- Brooks-Pollock, E. and Keeling, M., 2009. Herd size and bovine tuberculosis persistence in cattle farms in Great Britain. *Preventive veterinary medicine*, 92(4), pp.360-365.
- Brooks-Pollock, E., Conlan, A.J., Mitchell, A.P., Blackwell, R., McKinley, T.J. and Wood, J.L., 2013. Age-dependent patterns of bovine tuberculosis in cattle. *Veterinary research*, 44(1), p.97.

- Broughan, J.M., Maye, D., Carmody, P., Brunton, L.A., Ashton, A., Wint, W., Alexander, N., Naylor, R., Ward, K., Goodchild, A.V. and Hinchliffe, S., 2016. Farm characteristics and farmer perceptions associated with bovine tuberculosis incidents in areas of emerging endemic spread. *Preventive veterinary medicine*, 129, pp.88-98.
- Brown, E., Marshall, A.H., Mitchell, H.J. and Byrne, A.W., 2019. Cattle movements in Northern Ireland form a robust network: implications for disease management. *Preventive veterinary medicine*, 170, p.104740.
- Byrne, A.W., Barrett, D., Breslin, P., Madden, J.M., O'Keeffe, J. and Ryan, E., 2020a. Post-mortem surveillance of bovine tuberculosis in Ireland: herd-level variation in the probability of herds disclosed with lesions at routine slaughter to have skin test reactors at follow-up test. *Veterinary Research Communications*, pp.1-6.
- Byrne, A.W., Barrett, D., Breslin, P., Madden, J.M., O'Keeffe, J. and Ryan, E., 2020b. Future risk of bovine tuberculosis breakdown: a dominance analysis approach. *BioRxiv*.
- Cawley, A., Cronin, A. (2019). Spending Review 2019 Animal Health: TB Eradication. Irish Government Economic and Evaluation Service (IGEES), Government of Ireland. <http://budget.gov.ie/Budgets/2020/Documents/Budget/Animal%20Health%20-%20TB%20Eradication.pdf>
- Clarke, A., Maher, J., Byrne, A.W., More, S.J., Horan, M. and Barrett, D., 2020. Development and application of a prioritisation tool for animal health surveillance activities in Ireland. *Frontiers in Veterinary Science*, 5, In review.
- Clegg, T.A., Good, M., Hayes, M., Duignan, A., McGrath, G. and More, S.J., 2018. Trends and predictors of large tuberculosis episodes in cattle herds in Ireland. *Frontiers in Veterinary Science*, 5, p.86.
- De Garine-Wichatitsky, M., Caron, A., Kock, R., Tschopp, R., Munyeme, M., Hofmeyr, M. and Michel, A., 2013. A review of bovine tuberculosis at the wildlife–livestock–human interface in sub-Saharan Africa. *Epidemiology & Infection*, 141(7), pp.1342-1356.
- De Kantor, I.N. and Ritacco, V., 2006. An update on bovine tuberculosis programmes in Latin American and Caribbean countries. *Veterinary microbiology*, 112(2-4), pp.111-118.
- Doyle, L.P., Courcier, E.A., Gordon, A.W., O'Hagan, M.J.H. and Menzies, F.D., 2016. Bovine tuberculosis in Northern Ireland: Risk factors associated with duration and recurrence of chronic herd breakdowns. *Preventive Veterinary Medicine*, 131, pp.1-7.
- Gilbert, M., Mitchell, A., Bourn, D., Mawdsley, J., Clifton-Hadley, R. and Wint, W., 2005. Cattle movements and bovine tuberculosis in Great Britain. *Nature*, 435(7041), pp.491-496.
- Green, L.E. and Cornell, S.J., 2005. Investigations of cattle herd breakdowns with bovine tuberculosis in four counties of England and Wales using VETNET data. *Preventive veterinary medicine*, 70(3-4), pp.293-311.
- Griffin, J.M., Hahesy, T., Lynch, K., Salman, M.D., McCarthy, J. and Hurley, T., 1993. The association of cattle husbandry practices, environmental factors and farmer characteristics with the occurrence of chronic bovine tuberculosis in dairy herds in the Republic of Ireland. *Preventive Veterinary Medicine*, 17(3-4), pp.145-160.
- Griffin, J.M., Martin, S.W., Thorburn, M.A., Eves, J.A. and Hammond, R.F., 1996. A case-control study on the association of selected risk factors with the occurrence of bovine tuberculosis in the Republic of Ireland. *Preventive Veterinary Medicine*, 27(1-2), pp.75-87.
- Houtsma, E., Clegg, T.A., Good, M. and More, S.J., 2018. Further improvement in the control of bovine tuberculosis recurrence in Ireland. *The Veterinary Record*, 183(20), p.622.
- Jaki, T., Kim, M., Lamont, A., George, M., Chang, C., Feaster, D. and Van Horn, M.L., 2019. The effects of sample size on the estimation of regression mixture models. *Educational and Psychological Measurement*, 79(2), pp.358-384.
- Johnston, W.T., Gettinby, G., Cox, D.R., Donnelly, C.A., Bourne, J., Clifton-Hadley, R., Le Fevre, A.M., McInerney, J.P., Mitchell, A., Morrison, W.I. and Woodroffe, R., 2005. Herd-level risk factors associated with tuberculosis breakdowns among cattle herds in England before the 2001 foot-and-mouth disease epidemic. *Biology letters*, 1(1), pp.53-56.
- Karolemeas, K., McKinley, T.J., Clifton-Hadley, R.S., Goodchild, A.V., Mitchell, A., Johnston, W.T., Conlan, A.J.K., Donnelly, C.A. and Wood, J.L.N., 2010. Predicting prolonged bovine tuberculosis breakdowns in Great Britain as an aid to control. *Preventive Veterinary Medicine*, 97(3-4), pp.183-190.
- Kelly, P., Shalloo, L., Wallace, M. and Dillon, P., 2020. The Irish dairy industry—Recent history and strategy, current state and future challenges. *International Journal of Dairy Technology*, 73(2), pp.309-323.
- Läpple, D. and Sirr, G., 2019. Dairy intensification and quota abolition: A comparative study of production in Ireland and the Netherlands. *EuroChoices*, 18(3), pp.26-32.
- Magnussen, S., Næsset, E. and Gobakken, T., 2019. An application niche for finite mixture models in forest resource surveys. *Canadian Journal of Forest Research*, 49(11), pp.1453-1462.

- Manchia, M., Zai, C.C., Squassina, A., Vincent, J.B., De Luca, V. and Kennedy, J.L., 2010. Mixture regression analysis on age at onset in bipolar disorder patients: investigation of the role of serotonergic genes. *European Neuropsychopharmacology*, 20(9), pp.663-670.
- McLachlan, G.J. and McGiffin, D.C., 1994. On the role of finite mixture models in survival analysis. *Statistical methods in medical research*, 3(3), pp.211-226.
- McLachlan, G.J. and Peel, D., 2004. *Finite mixture models*. John Wiley & Sons.
- Mill, A.C., Rushton, S.P., Shirley, M.D.F., Murray, A.W.A., Smith, G.C., Delahay, R.J. and McDona ld, R.A., 2012. Farm-scale risk factors for bovine tuberculosis incidence in cattle herds during the Randomized Badger Culling Trial. *Epidemiology & Infection*, 140(2), pp.219-230.
- Milne MG, Graham J, Allen AR, Lahuerta-Marin A, McCormick C, Presho E, Skuce RA, Byrne AW (2018) Herd characteristics, wildlife risk and bacterial strain genotypes in persistent breakdowns of bovine tuberculosis in northern Irish cattle herds. Conference: Society for Veterinary Epidemiology and Preventive Medicine.
- Milne, G., Allen, A., Graham, J., Lahuerta-Marin, A., McCormick, C., Presho, E., Reid, N., Skuce, R. and Byrne, A.W., 2020. Bovine tuberculosis breakdown duration in cattle herds: an investigation of herd, host, pathogen and wildlife risk factors. *PeerJ*, 8, p.e8319.
- Milne, G.M., Graham, J., Allen, A., Lahuerta-Marin, A., McCormick, C., Presho, E., Skuce, R. and Byrne, A.W., 2019a. Spatiotemporal analysis of prolonged and recurrent bovine tuberculosis breakdowns in Northern Irish cattle herds reveals a new infection hotspot. *Spatial and spatio-temporal epidemiology*, 28, pp.33-42.
- Milne, M.G., Graham, J., Allen, A., McCormick, C., Presho, E., Skuce, R. and Byrne, A.W., 2019b. Variation in Mycobacterium bovis genetic richness suggests that inwards cattle movements are a more important source of infection in beef herds than in dairy herds. *BMC Microbiology*, 19(1), p.154.
- Moran, J.L., Solomon, P.J., ANZICS Centre for Outcome and Resource Evaluation (CORE) of the Australian and New Zealand Intensive Care Society (ANZICS, 2012. A review of statistical estimators for risk-adjusted length of stay: analysis of the Australian and New Zealand intensive care adult patient data -base, 2008–2009. *BMC medical research methodology*, 12(1), p.68.
- More, S.J., 2019. Perspectives from the science-policy interface in animal health and welfare. *Frontiers in Veterinary Science*, 6, p.382.
- Nowrouzi, B., Kamhi, R., Hu, J., Kennedy, J.L., Matmari, M. and De Luca, V., 2015. Age at onset mixture analysis and systematic comparison in schizophrenia spectrum disorders: Is the onset heterogeneity dependent on heterogeneous diagnosis?. *Schizophrenia Research*, 164(1-3), pp.83-91.
- Olea-Popelka, F.J., White, P.W., Collins, J.D., O'Keeffe, J., Kelton, D.F. and Martin, S.W., 2004. Breakdown severity during a bovine tuberculosis episode as a predictor of future herd breakdowns in Ireland. *Preventive veterinary medicine*, 63(3-4), pp.163-172.
- Reilly, L.A. and Courtenay, O., 2007. Husbandry practices, badger sett density and habitat composition as risk factors for transient and persistent bovine tuberculosis on UK cattle farms. *Preventive veterinary medicine*, 80(2-3), pp.129-142.
- Royston, P., 2001. The lognormal distribution as a model for survival time in cancer, with an emphasis on prognostic factors. *Statistica Neerlandica*, 55(1), pp.89-104.
- Skuce, R.A., Allen, A.R. and McDowell, S.W., 2012. Herd-level risk factors for bovine tuberculosis: a literature review. *Veterinary Medicine International*, 2012.
- Van Horn, M.L., Smith, J., Fagan, A.A., Jaki, T., Feaster, D.J., Masyn, K., Hawkins, J.D. and Howe, G., 2012. Not quite normal: Consequences of violating the assumption of normality in regression mixture models. *Structural equation modeling: a multidisciplinary journal*, 19(2), pp.227-249.
- Vial, F., Johnston, W.T. and Donnelly, C.A., 2011. Local cattle and badger populations affect the risk of confirmed tuberculosis in British cattle herds. *PLoS One*, 6(3), p.e18058.
- Wolfe, D.M., Berke, O., Kelton, D.F., White, P.W., More, S.J., O'Keeffe, J. and Martin, S.W., 2010. From explanation to prediction: a model for recurrent bovine tuberculosis in Irish cattle herds. *Preventive veterinary medicine*, 94(3-4), pp.170-177.