The Veterinary Journal 238 (2018) 41–48

Contents lists available at ScienceDirect

The Veterinary Journal

journal homepage: www.elsevier.com/locate/tvjl

Invited Review

Use of statistical modelling to investigate the pathogenesis of claw horn disruption lesions in dairy cattle

L.V. Randall*, M.J. Green, J.N. Huxley

University of Nottingham, School of Veterinary Medicine and Science, Sutton Bonington Campus, Sutton Bonington, Leicestershire, LE12 5RD, UK

ARTICLE INFO

Article history:
Accepted 13 July 2018

Keywords:
Claw horn disruption lesions
Dairy cattle
Pathogenesis
Risk factors
Statistical modelling

ABSTRACT

Claw horn disruption lesions (CHDLs) in dairy cattle account for a large proportion of lameness. The aim of this review is to provide an update on the evidence surrounding the pathogenesis of CHDLs, in the context of how statistical modelling has contributed to the validity of available evidence and current thinking. Historically, 'subclinical laminitis' has often been used to describe the commonly accepted underlying pathology associated with these lesions, however progress in understanding the aetiopathogenesis of CHDLs and a lack of clear evidence to support the traditional laminitis hypothesis, means use of this terminology has been challenged. With advancements in statistical modelling capabilities within the veterinary field, the multifactorial and complex nature of CHDLs can be more fully explored. This has led to an increased understanding of environmental and animal-based risk factors and their role in the pathogenesis of CHDLs, as well as highlighting future research areas. There is still a need for further research using intervention studies to demonstrate causality for identified risk factors to date, as well as quantifying the impact of these risk factors at the population level. Some important considerations when using and interpreting statistical models in lameness research are discussed with a critical assessment of the key statistical issues in published research investigating the pathogenesis of CHDLs.

Introduction

Lesions of claw horn disruption, primarily sole ulceration, sole haemorrhage and white line disease (haemorrhage and separation), account for a large proportion of lameness in dairy cattle (Manske et al., 2002; Sogstad et al., 2005). Sole ulceration was first described as Rusterholz disease in the 1920s and since then studies investigating the aetiopathogenesis of claw horn disruption lesions (CHDLs) have led to the identification of a wide range of risk factors (Hirst et al., 2002a). These can be broadly generalised into two categories; animal based risk factors (i.e. internal factors originating within the animal; predominantly factors related to the structure and function of the claw) and environmental risk factors (i.e. external factors which operate beyond the individual; predominantly factors which directly or indirectly increase the pressure on the hoof capsule).

One of the earliest hypotheses describing the pathogenesis of CHDLs was related to the occurrence of 'laminitis' (Nilsson, 1963); a proposal which was principally based on the symptoms and pathogenesis described for equine laminitis at that time. Studies investigating laminitis became predominant throughout the lameness literature in the subsequent years. However limitations in statistical analysis and techniques used were evident at this time, such that the multifactorial nature of CHDLs could not be fully explored. More recently, with the advancement and availability of statistical software and computing power, statistical modelling has been used to further our understanding of the complex nature of CHDLs in cattle by conducting multivariable analyses. Risk factors related to the environment have been investigated alongside animal-based risk factors and alternative hypotheses that may explain associations previously identified in the literature have been suggested.

To evaluate how statistical modelling has been used to investigate the pathogenesis of CHDLs, consideration must extend from the study design that generates data through to interpretation of results generated by models. All models are a simplification of reality, therefore both transparency and accuracy regarding model assumptions and reporting are important to enable understanding of the disease epidemiology (Huppert and Katriel, 2013). This review aims to provide an update on the current evidence surrounding the pathogenesis of CHDLs from the perspective of how statistical modelling has contributed to the evolution of our understanding of CHDLs and associated risk factors, since the laminitis hypothesis was proposed in the 1960s.

* Corresponding author.
E-mail address: Laura.Randall@nottingham.ac.uk (L.V. Randall).

https://doi.org/10.1016/j.tvjl.2018.07.002
1090-0233/

Since the scope of the paper focuses on the use of statistical modelling, a comprehensive review of all the available literature on the pathogenesis of CHDLs is not provided. For reviews on the lameness literature readers are referred to Hirst et al. (2002a) and Potterton et al. (2012).

General considerations for statistical modelling in lameness research

Statistical methods can be used to analyse relationships between measurements on groups of animals and statistical models to provide a mathematical description of these relationships (Dobson, 1983). The process of statistical modelling can be described in three stages (Dobson, 1983); (1) specifying the equations and distributions that describe the primary features of the outcome, (2) estimating parameters (e.g. risk factor that is being investigated and confounding factors) and (3) making inferences. No model is perfect, but in order to evaluate how well a model describes the data (and therefore the validity of reported findings), there are a number of areas to consider from study design through to making inferences. This section describes some general considerations that are specific to lameness research with some examples, whilst further critical assessment of how statistical modelling has contributed to current evidence and thinking are discussed in subsequent sections.

Study design

Within the totality of peer-reviewed lameness literature, observational and cross-sectional studies are the most commonly reported study design (Hirst et al., 2002a; Potterton et al., 2012). Whilst variables that are found to be significantly associated with a lameness outcome may be causally linked, in general, results from these types of study provide weak evidence for causality. Dohoo et al. (2003) outlines a set of criteria for demonstrating causality including; time sequence, plausibility and experimental evidence. Cohort studies and randomised controlled trials (RCTs) can provide stronger evidence to support causality, yet they are under-represented in studies investigating the pathogenesis of CHDLs. This is likely due to the high resource demands with this type of study design, including cost and time and the accompanying relative lack of funding for lameness research.

Sample sizes and sampling procedures (e.g. randomization) are an important aspect of study design, however they are very commonly under-reported in the lameness literature. This is particularly evident in experimental research investigating laminitis where often relatively small numbers of animals were included e.g. Danscher et al. (2010) where group sizes were less than 10 animals. Sample size calculations indicate the number of animals required in each study group to demonstrate a significant difference; without this information it is not possible to interpret the relevance of negative findings with any confidence. However, sample size calculations for correlated data (e.g. repeated measures), such as that commonly encountered when investigating lameness, are not straightforward and bespoke software is often required (for example, GLIMMPSE (Kreidler et al., 2013)) (Liu and Liang, 1997; Guo et al., 2013).

Data analysis and statistical modelling techniques

Besides study design, statistical modelling can help to control for confounding factors in lameness studies investigating CHDLs. Potential confounders therefore need to be considered, identified and measured for data to be analysed appropriately. Regression analysis is a commonly used statistical technique that enables a number of variables to be incorporated into the model simultaneously (e.g. days in milk, milk yield, body condition score (BCS)); 'multivariable' models therefore enable confounding factors to be controlled for. Univariable statistics on the other hand do not allow for control of confounding factors. This is a major limitation for many of the early studies investigating laminitis, which were primarily observational studies and/or conducted under field conditions where confounding factors will almost inevitably exist (e.g. Bazeley and Pinsent (1984); Manson and Leaver (1988b)). It was more recently (post-2000) that the use of multivariable statistics has increasingly been used to explore the multifactorial nature of CHDLs, helping to progress understanding of animal-based risk factors, such as BCS (Green et al., 2014; Randall et al., 2015; Newsome et al., 2017b), as well as environmental risk factors such as alley and track surface (Barker et al., 2009; Burow et al., 2014).

Data gathered for lameness research commonly have a hierarchical (clustered) structure, for example, repeated measures of lameness (e.g. lameness score or treatment events (level 1) within a cow (level 2) within a farm (level 3)). Where this occurs, similarities may exist between the units at each level such that outcomes may be correlated. It is important that statistical modelling recognises and accounts for this to avoid biased parameter estimates (Woodhouse and Goldstein, 1988; Rabash et al., 2009). Modelling techniques that account for such correlations (e.g. mixed effects or multilevel) should therefore be used in any lameness studies with this data structure, however until more recently when software such as SAS (e.g. GLIMMIX procedure) or MLWin (Rabash et al., 2009) became more widely available, this has not been the case. In much of the early lameness research this was a major limitation, contributing to the publication of a large body of work providing weak evidence in support of the laminitis hypothesis.

Reporting of model assumptions and evaluation of model fit (i.e. how well the model fits the data) is a critical step and is another aspect to modelling that is under-reported in the lameness literature investigating CHDLs e.g. Haskell et al. (2006), Vanegas et al. (2006) and Bergsten et al. (2015). Statistical models that have a poor fit to the data can lead to erroneous conclusions being drawn. Key principles for evaluating model fit include; (1) checking model assumptions, (2) assessment of model fit by comparison of model predictions with observed data and (3) using cross-validation to determine likely generalisability of models (ideally a new data set although this is rarely possible). Appropriate methods for assessing model fit in mixed effects models have previously been described elsewhere (Gelman et al., 1996; Green et al., 2009).

A particularly pertinent issue in the lameness literature is the interdependence between risk factors and outcomes. Identifying the direction of causality can therefore be complex; for example lying times impact on lameness, which in itself impacts on lying times. Complete historical data and appropriate statistical modelling is particularly important in longitudinal studies in order to understand these temporal associations. This is relevant for studies investigating both animal-based (e.g. BCS or body weight (BW)) and environmental risk factors (e.g. those related to lying times) for CHDLs.

Reporting results and interpretation

Odds ratios (OR) estimated from logistic regression models (where the outcome is binary) are very commonly reported in studies investigating the pathogenesis of CHDL. Odds ratios and relative risk (RR) are similar, however when the disease event is common (e.g. lameness incidence rates of between 20 and 80 cases per hundred cows are not uncommon), OR may be substantially different to the RR and care is needed when interpreting results.