Dairy herd mastitis and reproduction: Using simulation to aid interpretation of results from discrete time survival analysis

Christopher D. Hudson [a, *], Andrew J. Bradley [a, b], James E. Breen [a, b], Martin J. Green [a]

[a] *School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington, LE12 5RD, UK*
[b] *Quality Milk Management Services Ltd, Cedar Barn, Easton, Wells, Somerset, BA5 1DU, UK*

* Corresponding author: Tel.: +44 115 9516182
E-mail address: chris.hudson@nottingham.ac.uk (C.D. Hudson)

Abstract

Probabilistic sensitivity analysis (PSA) is a simulation-based technique for evaluating the relative importance of different inputs to a complex process model. It is commonly employed in decision analysis and for evaluation of the potential impact of uncertainty in research findings on clinical practice, but has a wide variety of other possible applications. In this example, it was used to evaluate the association between herd-level udder health and reproductive performance in dairy herds.

Although several recent studies have found relatively large associations between mastitis and fertility at the level of individual inseminations or lactations, the current study demonstrated that herd-level intramammary infection status is highly unlikely to have a clinically significant impact on the overall reproductive performance of a dairy herd under typical conditions. For example, a large increase in incidence rate of clinical mastitis (from 92 to 131 cases per 100 cows per year) would be expected to increase a herd's modified FERTEX score (a cost-based measure of overall reproductive performance) by just £4.50[1] per cow per year. The herd's background level of submission rate (proportion of eligible cows served every 21 days) and pregnancy risk (proportion of inseminations leading to a pregnancy) correlated strongly with overall reproductive performance and explained a large proportion of the between-herd variation in performance.

PSA proved to be a highly useful technique to aid understanding of results from a complex statistical model, and has great potential for a wide variety of applications within the field of veterinary science.

[1] £1 = approx. US$1.61, €1.26 at 17 October 2014

Keywords: Bayesian, Dairy cow, Fertility performance, Mastitis, Probabilistic sensitivity analysis

Introduction

As the volume and reliability of data routinely recorded by dairy herds grows, the potential for large-scale epidemiological studies in the field increases. These often require sophisticated analytical techniques, which can make interpretation of their practical consequences challenging. In many cases, research yields important information on a particular aspect of a biological system, but it can be difficult to see the results in the context of the system as a whole. For example, the reproductive performance of a dairy herd is a complex, multi-factorial system and, although detailed knowledge exists about many specific elements of this system, it can be difficult to evaluate how such knowledge fits together to determine the overall reproductive outcome. For instance, there have been a number of recent publications demonstrating associations between a cow's udder health and the probability of conceiving to a specific insemination or during a given period of lactation (Hertl et al., 2010; Lavon et al., 2011; Hudson et al., 2012), but the likely importance of this at the herd level is unclear. For decision makers, it remains difficult to evaluate the potential improvement in a herd's reproductive performance that might be expected if udder health on the farm were improved.

A prominent technique for studying the relative importance of different inputs into a complex system is known as probabilistic sensitivity analysis (PSA). PSA is a stochastic, simulation-based approach, whereby the input values for a system are drawn from pre-defined probability distributions. At each iteration of the simulation, a value for each input is drawn at random from the relevant distribution. A mathematical model is then used to convert the inputs into one or more output values, often through complex inter-relationships, and results are stored for that iteration. The distribution of output values across the iterations, and the correlations between specific inputs and any output of interest can then be analysed,

providing a way to evaluate the relative extent to which different model inputs affect outcome.

Although PSA is perhaps most commonly applied to cost-effectiveness analysis in medicine (Spiegel et al., 2003; Anderson et al., 2006; Gillies et al., 2008), it has been used in a variety of alternative contexts (Steinbach et al., 2012) and has huge potential in the evaluation of the likely effectiveness of population-level interventions and in integrating multiple sources of research knowledge. PSA allows a degree of model complexity limited only by computational power and provides a robust way of evaluating the relative importance of different inputs to a system even where such inputs are inter-correlated. Despite these advantages, use of PSA as a tool to understand the action of complex biological systems is still relatively uncommon, and reports of such approaches in veterinary science are still rare (Detilleux, 2004; Heller et al., 2011).

In this study, PSA was used to evaluate the relative importance of different model inputs where minimal assumptions were made about the distribution of input parameters (i.e. under conditions of extreme uncertainty): that is, all values within a specified range were equally likely to be drawn at each iteration. We aimed to evaluate the likely scope for change in a herd's reproductive performance which could result from an improvement in intramammary infection status, relative to the other factors which affect fertility.

Materials and methods

Discrete time survival model

The study was based on a statistical model previously developed to describe reproductive performance in dairy cows by predicting the probability that a given cow would

become pregnant in each consecutive 2-day risk period throughout lactation. Explanatory variables significantly associated with this outcome were used as the input parameters for the simulation model described here. This statistical model has been described in detail in a previous publication (Hudson et al., 2012), but is summarised in Appendix A.

Distributions of simulation input variables

The distributions of the simulation input parameters are described in Table 1. Independent uniform distributions were selected for all herd-level inputs, covering ranges considered likely to encompass true values for the vast majority of UK herds. Although these distributions were not intended to represent the true 'real world' distributions of the inputs, ranges were selected so that evaluation was carried out across the full range of plausible herd-level scenarios. These were treated as equally likely by assigning a uniform probability across the range for each input parameter.

The input parameters for each lactation, and for each risk period within the lactation, were mostly dependent on herd level inputs, so were drawn from appropriate distributions based on the relevant herd level parameter (Table 1). The possibility that correlations between the input parameters would affect the outcome of the simulation was also explored (for details, see Appendix A).

Simulation model

The structure of the simulation model is represented diagrammatically in Fig. 1. Simulation was carried out in Excel 2010 (Microsoft), using Visual Basic for Applications (Microsoft) for process control. A total of 50,000 herds were simulated, with each one consisting of 200 lactations.

The first step in simulating a herd was to draw the herd level input parameters from their distributions before simulating the first lactation in the herd (again, beginning by drawing the lactation level inputs from relevant distributions). Next, a simulated udder health history was generated for the lactation (Fig. 2; see Appendix A for detail). The logistic regression model from Hudson et al. (2012; also described in Appendix A) was then used to calculate the probability of pregnancy occurring during each 2-day risk period of the lactation (based on the input parameters for that herd, lactation and risk period). This probability was then adjusted to account for additional marginal (i.e. unexplained by model input parameters) variation in the herd's submission rate (proportion of eligible cows served every 21 days) and pregnancy risk (proportion of inseminations leading to a pregnancy).

A binary outcome for pregnancy in each 2-day risk period was then drawn from a binomial distribution based on this adjusted probability, with repeated risk periods simulated until either pregnancy or 300 days in milk (DIM). The reproductive outcome of the lactation was recorded using two variables, namely, a binary outcome representing whether the cow reached 300 DIM without becoming pregnant, and, if the cow did become pregnant, the number of DIM at which pregnancy occurred. This information was stored along with the input parameters for the lactation, and simulation of the next lactation begun.

The process was repeated until the 200 lactations making up the herd were complete, at which point the mean number of DIM to pregnancy (i.e. calving to conception interval) and the proportion of lactations where the cow reached 300 DIM without becoming pregnant were calculated over the herd and stored, along with the herd input parameters. These two measures were combined to produce a single outcome using a modification of the 'FERTEX'

140 score (Esslemont and Kossaibati, 2002) (mFX), described in full in Appendix A. Simulation
141 of the next herd was then begun.
142
143 *Analysis of results*
144 Summary data for each of the 50,000 simulated herds were exported to R 2.14.2 (R
145 Core Development Team, 2010) for analysis. The associations between each herd-level input
146 parameter and the outcome (mFX score) were initially explored using high-density
147 scatterplots. High-density (or 'heatmap') scatterplots are bivariate density plots where the
148 density of points at any given location is represented by colour darkness; these were required
149 as there were a very large number of points (i.e. simulated herds) to be represented. As the
150 mFX scores were strongly positively skewed (as expected with a cost-based outcome),
151 Spearman rank correlation coefficients were calculated for the relationships between mFX
152 score and each input.
153
154 Multiple regression, with the natural logarithm of herd mFX score as the outcome
155 variable, was used to partition variance in mFX score between the herd input parameters, and
156 to predict the effect of changes in each individual parameter on herd mFX score. In order to
157 represent these results graphically as a tornado plot, the predicted change in mFX score was
158 calculated where each input parameter in turn was increased from the median value of its
159 input distribution by a value representing 25% of the range of the distribution while the other
160 inputs were held at their median values. This allowed evaluation of the change in outcome
161 (mFX score) when each input parameter was altered by a comparable amount, allowing
162 visualisation of relative effect size.

Results

Univariate analysis

High density scatterplots showing the associations between each herd-level input parameter and the herd mFX score (with higher mFX scores indicating poorer overall performance), along with the Spearman rank correlation coefficient (r_s) for each relationship are shown in Fig. 3. The association between herd submission rate and mFX score was the most striking, with a clear 'funnelling' of points in the bottom right hand corner of the graph, indicating that herds with high submission rates (especially over 50%) had a much narrower range of mFX scores, with a much stronger concentration around the lower mFX scores (i.e. better reproductive performance). The high-density scatterplots showing relationship between the udder-health-related input parameters and mFX score showed no correlations, with point clouds assuming a square appearance and no evident trend in the line of highest point density.

Multiple regression analysis

The results of variance partition by regression analysis are shown in Table 2. Each line of the table shows the proportion of variation in mFX score explained by each input parameter, after accounting for the variation explained by the other input parameters. It is clear that submission rate (42.9% of total variance) and pregnancy risk (35.2% of total variance) collectively account for the vast majority of variance in the outcome.

The predicted effects of changes in inputs are represented graphically as a tornado plot in Fig. 4. Changing submission or pregnancy risk was predicted to have a large impact on overall reproductive performance, with a move from median (45%) to upper quartile (62.5%) submission rate predicted to generate a saving of more than £85 per cow per year: Cost per additional day on calving index and average 305-day adjusted milk yield were

188 associated with smaller changes in mFX score, and cost per cull predicted to lead to a slightly
189 smaller change again. Udder-health-related inputs were predicted to have little impact on
190 overall reproductive performance.
191
192 The low degree of association between udder health parameters and herd reproductive
193 performance is demonstrated further in Fig. 5 – Figs. 5a and b show the distributions (as
194 kernel density plots) of mFX scores for herds with extremely high or low values for incidence
195 rates of clinical mastitis or proportion of individual cow somatic cell count (ICSCC)
196 recordings >200k, respectively. The two lines on each figure follow a very similar shape,
197 demonstrating that herds at either extreme of the distribution for udder health parameters had
198 very similar ranges of reproductive performance. By contrast, Fig. 5c shows the distributions
199 of mFX scores for herds with extremely high and extremely low submission rates; herds with
200 high submission rates have a much tighter distribution of mFX scores centred on a much
201 lower mFX score compared to low submission rate herds.
202
203 The analysis was repeated on the subsets of simulated herds with very high marginal
204 submission rates and pregnancy risks (>70% and 45%, respectively) and very low marginal
205 submission rates and pregnancy risks (< 20% and 25%, respectively). This revealed very
206 similar results, with very little clear relationship between udder health parameters and herd
207 reproductive performance under either scenario (i.e. in herds with exceptionally good or poor
208 'background' performance).
209

Discussion

Recent work has demonstrated that clinical mastitis around the time of insemination is associated with a reduction in the probability of pregnancy to the insemination of between 20 and 80% (Hertl et al., 2010; Hudson et al., 2012), and that elevated ICSCC can be associated with reductions in the order of 20% (Lavon et al., 2011; Hudson et al., 2012). However, although these effect sizes intuitively appear quite large and are broadly consistent with earlier work in the area (Loeffler et al., 1999; Schrick et al., 2001; Pinedo et al., 2009), interpreting their likely impact at herd level has been difficult owing to the large number of other factors that influence the relationship between mastitis and reproduction (for example, the frequency and distribution of clinical mastitis cases and elevations of ICSCC throughout lactation). Specifically, these results did not give farmers or veterinary surgeons any indication of the potential to improve a herd's reproduction by maximising udder health.

Here, development of a simulation model and its use within a PSA framework have revealed that improvements in udder health at herd level are highly unlikely to lead to useful improvement in herd fertility performance under the vast majority of plausible scenarios. Therefore, given the variability in udder health performance typically observed in UK dairy herds (represented by the ranges chosen for the distributions of the input parameters), it is highly unlikely that improving a herd's udder health (either in terms of clinical mastitis or somatic cell count) would lead to a detectable improvement in the reproductive performance of the herd. The study also confirmed that the marginal effects of submission rate and pregnancy risk (after accounting for effects of other model inputs, such as milk yield) are key drivers of performance, and gave an indication of the potential room for investment in these areas.

Use of stochastic modelling (and associated techniques such as PSA) is becoming increasingly commonplace in a variety of areas. Essentially, such models have two main applications. Firstly, they can be used in a research setting to evaluate the likely importance of different model inputs across a variety of possible scenarios. Results of such research can then be used to inform clinical guidance, as well as prioritising promotion of existing knowledge and allocation of resources towards future research. Clinical decision making in human medicine presents an excellent example here, with PSA widely adopted for cost-effectiveness studies informing blanket clinical guidelines (Andronis et al., 2009).

Secondly, stochastic modelling can be used on a case-by-case basis, whereby simulation using a model can be used to evaluate likely outcomes for a specific real-life scenario under alternative potential strategies or interventions. Risk management in business (especially the financial sector) presents perhaps the best example of this process: for example, use of such tools is extremely common for evaluation of alternative investment opportunities. It is easy to see excellent uses for both of these approaches in clinical veterinary medicine (especially in farm animal practice, where decisions regarding potential interventions at herd level are common). Despite this, early efforts to develop a decision support tool for dairy herds along these lines (Sørensen et al., 1992) has not led to widespread uptake, and although there is increasing use of stochastic models in research they tend to be at a 'macro' or 'whole farm' level (Geary et al., 2012) rather than the 'micro' level described in this study; and use of PSA in the veterinary literature is still uncommon.

Recently, there has been more interest in both applications of stochastic modelling to herd-level management decisions in dairy farms, but it is often considered that such methods are too complex and cumbersome to be widely employed by farmers or their advisors

(Walster, 2012). However, the simulation model in this paper was deliberately developed in a software environment that would allow for development of customised decision support tools, based on the approach described, which could be widely distributed and used within the industry.

Whilst PSA is a robust and well established technique, a common criticism is that unjustified assumptions are made about parameter input distributions. In this case PSA was being used to evaluate dairy herd reproduction as a system and assess which input parameters are most able to perturb the system: effectively this represented simulating hypothetical herds across as wide a range of plausible situations as possible. This is the reason uniform distributions were used for the input parameters. Although these clearly do not reflect the distributions of the same parameters across real life herds, they allow the relative importance of each parameter to be evaluated across a wide variety of possible scenarios. The udder health inputs are a good example of this, with clinical mastitis and somatic cell count history through each lactation were simulated independently. In reality, these are both driven by an underlying latent variable (the true intramammary infection status through lactation), which is difficult to evaluate and therefore to simulate realistically. However, as their overall effects appear to be very small, this is not likely to have made a substantive difference to the results of this study. In this case, it also appeared that using independent input distributions did not lead to a different conclusion than that reached using the observed joint distributions from the original data (see Appendix A).

Conclusions

This study has found that the association between herd intramammary infection status (as measured by clinical mastitis and ICSCC) and herd-level reproductive performance is likely to be weak under the vast majority of plausible scenarios, despite the relatively large association sizes at lactation and service level revealed by previous work and used as model inputs. In this example, development of a stochastic model and PSA were found to be useful tools to aid understanding of dairy herd reproduction as a system. Importantly, this work has also provided a model structure that can be extended and built upon in future research.

Conflict of interest statement

None of the authors has any financial or personal relationships that could inappropriately influence or bias the content of the paper.

Appendix: Supplementary material

Supplementary data associated with this article can be found in the online version.

References

Anderson, K., Jacobson, J.S., Heitjan, D.F., Zivin, J.G., Hershman, D., Neugut, A.I., Grann, V.R., 2006. Cost-effectiveness of preventive strategies for women with a BRCA1 or a BRCA2 mutation. Annals of Internal Medicine 144, 397–406.

Andronis, L., Barton, P., Bryan, S., 2009. Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making. Health Technology Assessment 13, 1–45.

Detilleux, J.C., 2004. Javelin diagrams: applications in veterinary medical decision analysis. Veterinary Research 35, 617–624.

Esslemont, R.J., Kossaibati, M., 2002. DAISY Research Report No. 5: The costs of poor fertility and disease in UK dairy herds - Trends in DAISY herds over 10 seasons. University of Reading, UK.

309 Geary, U., Lopez-Villalobos, N., Begley, N., McCoy, F., O'Brien, B., O'Grady, L., Shalloo,
310 L., 2012. Estimating the effect of mastitis on the profitability of Irish dairy farms.
311 Journal of Dairy Science 95, 3662–3673.

312 Gillies, C.L., Lambert, P.C., Abrams, K.R., Sutton, A.J., Cooper, N.J., Hsu, R.T., Davies,
313 M.J., Khunti, K., 2008. Different strategies for screening and prevention of type 2
314 diabetes in adults: cost effectiveness analysis. British Medical Journal 336, 1180–
315 1191.

316 Green, M.J., Green, L.E., Medley, G.F., Schukken, Y.H., Bradley, A.J., 2002. Influence of
317 dry period bacterial intramammary infection on clinical mastitis in dairy cows.
318 Journal of Dairy Science 85, 2589–2599.

319 Heller, J., Innocent, G.T., Denwood, M., Reid, S.W.J., Kelly, L., Mellor, D.J., 2011.
320 Assessing the probability of acquisition of methicillin-resistant Staphylococcus aureus
321 (MRSA) in a dog using a nested stochastic simulation model and logistic regression
322 sensitivity analysis. Preventive Veterinary Medicine 99, 211–224.

323 Hertl, J.A., Grohn, Y.T., Leach, J.D.G., Bar, D., Bennett, G.J., Gonzalez, R.N., Rauch, B.J.,
324 Welcome, F.L., Tauer, L.W., Schukken, Y.H., 2010. Effects of clinical mastitis
325 caused by gram-positive and Gram-negative bacteria and other organisms on the
326 probability of conception in New York State Holstein dairy cows. Journal of Dairy
327 Science 93, 1551–1560.

328 Hudson, C.D., Bradley, A.J., Breen, J.E., Green, M.J., 2012. Associations between udder
329 health and reproductive performance in United Kingdom dairy cows. Journal of Dairy
330 Science 95, 3683–3697.

331 Lavon, Y., Ezra, E., Leitner, G., Wolfenson, D., 2011. Association of conception rate with
332 pattern and level of somatic cell count elevation relative to time of insemination in
333 dairy cows. Journal of Dairy Science 94, 4538–4545.

334 Loeffler, S.H., de Vries, M.J., Schukken, Y.H., 1999. The effects of time of disease
335 occurrence, milk yield, and body condition on fertility in dairy cows. Journal of Dairy
336 Science 82, 2589 – 2604.

337 Pinedo, P.J., Melendez, P., Villagomez-Cortes, J.A., Risco, C.A., 2009. Effect of high
338 somatic cell counts on reproductive performance of Chilean dairy cattle. Journal of
339 Dairy Science 92, 1575–1580.

340 R Core Development Team, 2010. R: A language and environment for statistical computing.
341 R Foundation for Statistical Computing, Vienna, Austria.

342 Schrick, F.N., Hockett, M.E., Saxton, A.M., Lewis, M.J., Dowlen, H.H., Oliver, S.P., 2001.
343 Influence of subclinical mastitis during early lactation on reproductive parameters.
344 Journal of Dairy Science 84, 1407–1412.

345 Sørensen, J.T., Kristensen, E.S., Thysen, I., 1992. A stochastic model simulating the dairy
346 herd on a PC. Agricultural Systems 39, 177–200.

347 Spiegel, B.M.R., Targownik, L., Dulai, G.S., Gralnek, I.M., 2003. The cost-effectiveness of
348 cyclooxygenase-2 selective inhibitors in the management of chronic arthritis. Annals
349 of Internal Medicine 138, 795–806.

350 Steinbach, R., Cairns, J., Grundy, C., Edwards, P., 2012. Cost benefit analysis of 20 mph
351 zones in London. Injury Prevention 19, 211–213.

352 Walster, C., 2012. Adding value for farmers through herd health plans. Veterinary Record
353 170, 524.

354

355 **Table 1**

356 Input parameters used at each level of simulation and distributions from which inputs were

357 drawn.

Input variable	Type	Input distribution
Herd level		
Submission rate (proportion of eligible cows inseminated every 21 days)	Continuous	Uniform (0.1, 0.8)
Pregnancy risk (proportion of inseminations leading to a pregnancy)	Continuous	Uniform (0.1, 0.6)
Herd average 305 day milk yield (kg)	Continuous	Uniform (3000, 12500)
Proportion of herd which are first lactation	Continuous	Uniform (0.1, 0.4)
Herd incidence rate of clinical mastitis (cases per cow-year of risk)	Continuous	Uniform (0.15, 1.7)
Proportion of clinical mastitis cases originating from dry period infection	Continuous	Uniform (0.1, 0.9)
Proportion of cows beginning lactation with ICSCC >200k	Continuous	Uniform (0.02, 0.4)
Proportion of cows moving from ICSCC <200k to >200k between milk recording test days	Continuous	Uniform (0.02, 0.25)
Proportion of cows moving from ICSCC >200k to <200k between milk recording test days	Continuous	Uniform (0.05, 0.45)
Cost per day of extension of calving index (£)	Continuous	Uniform (1.2, 4.2)
Cost per cow culled for failure to conceive (£)	Continuous	Uniform (550, 1750)
Lactation level		

Lactation number	Categorical (1, 2, 3, 4, >4)	Multinomial, based on proportion of herd in lactation 1
305 day milk yield (kg)	Continuous	Beta, centred on herd average with standard deviation 1.5k
Risk period level		
Season (quarter of year)	Categorical (1, 2, 3, 4)	Multinomial for season at calving
Occurrence of CM 15-28 days before risk period	Binary	Yes/No
Occurrence of CM 1-7 days before risk period	Binary	Yes/No
Occurrence of CM during risk period	Binary	Yes/No
Occurrence of CM 1-7 days after risk period	Binary	Yes/No
Occurrence of CM 8-14 days after risk period	Binary	Yes/No
Occurrence of CM 15-28 days after risk period	Binary	Yes/No
Occurrence of CM 29-42 days after risk period	Binary	Yes/No
Occurrence of CM 43-56 days after risk period	Binary	Yes/No
Occurrence of CM 57-70 days after risk period	Binary	Yes/No
ICSCC 1-30 days after risk period	Binary	(<=200k, >200k)

358 ICSCC, individual cow somatic cell count; CM, clinical mastitis

Table 2

Partition of variance in modified herd FERTEX score (mFX) between input parameters.

Input parameter	% variance explained
Submission rate	42.9%
Pregnancy risk	35.2%
305 day yield	7.4%
Incidence rate of CM	0.1%
% ICSCC recordings >200k	0.1%
% CM cases which are of dry period origin	<0.1%
% of herd in first lactation	<0.1%
Cost per day on calving index	5.5%
Cost per cull	1.3%
Total	92.5%

ICSCC, individual cow somatic cell count; CM, clinical mastitis

Figures

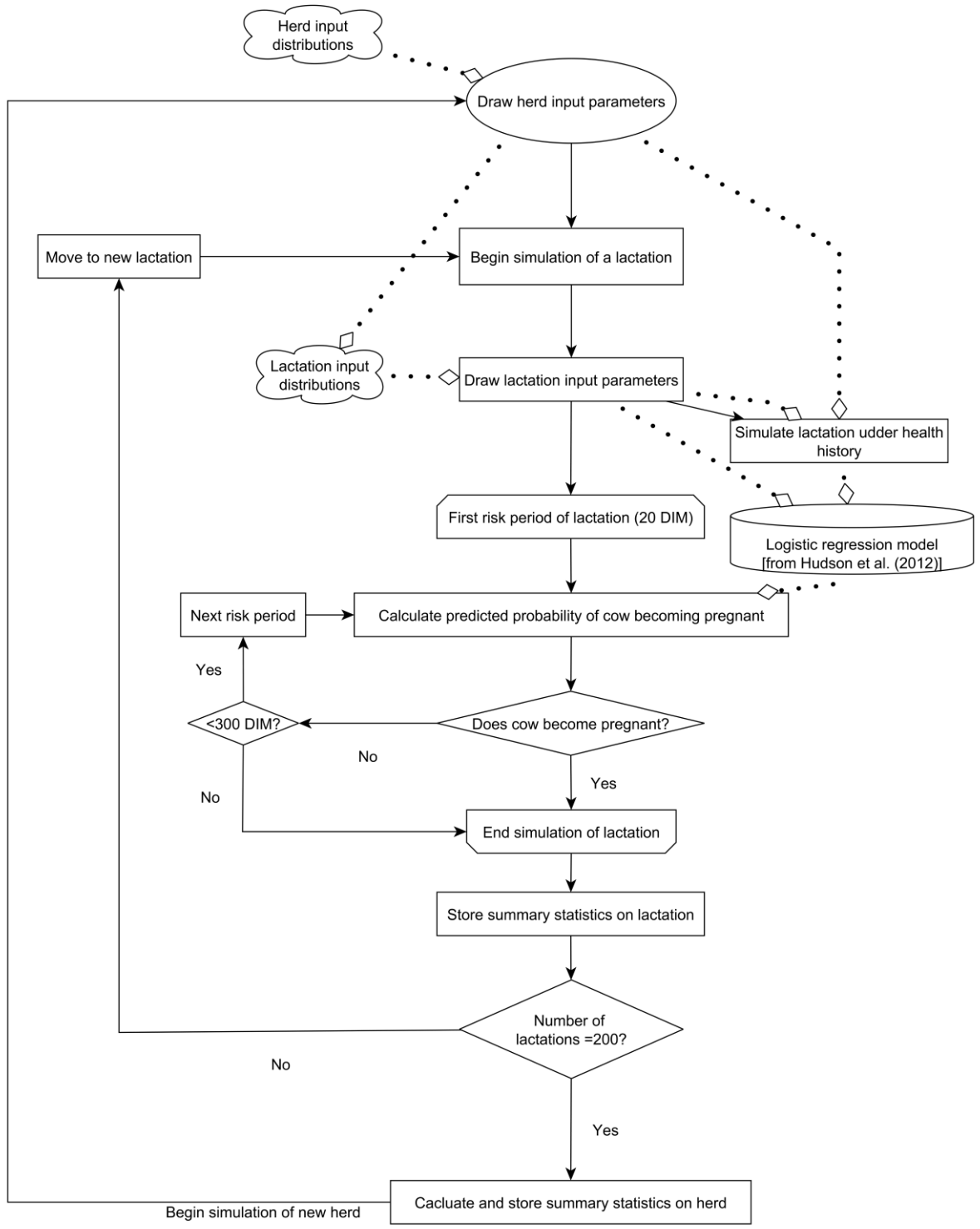
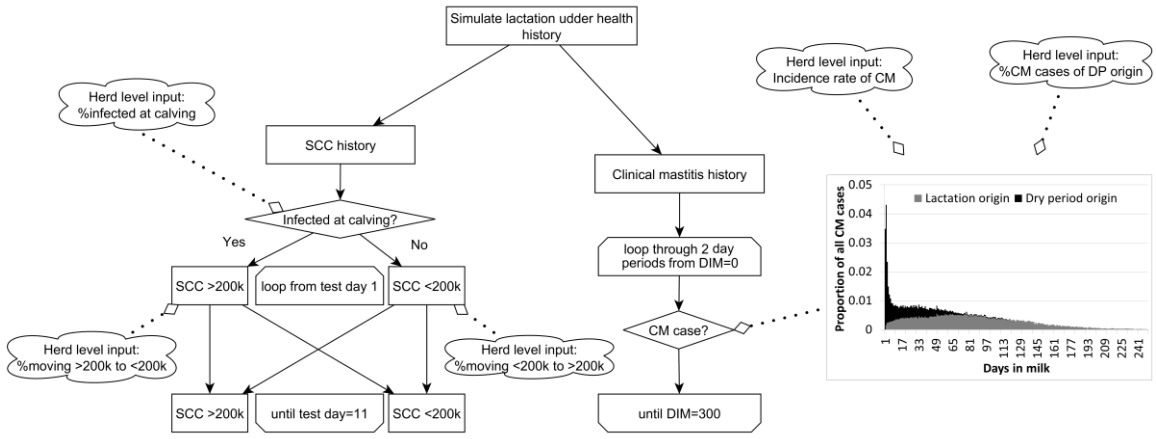


Fig. 1: Overview of the simulation model process. Solid black lines indicate process flow, and dotted lines indicate that information from the source of the line is used in the step of the process to which the line leads (denoted by a diamond).

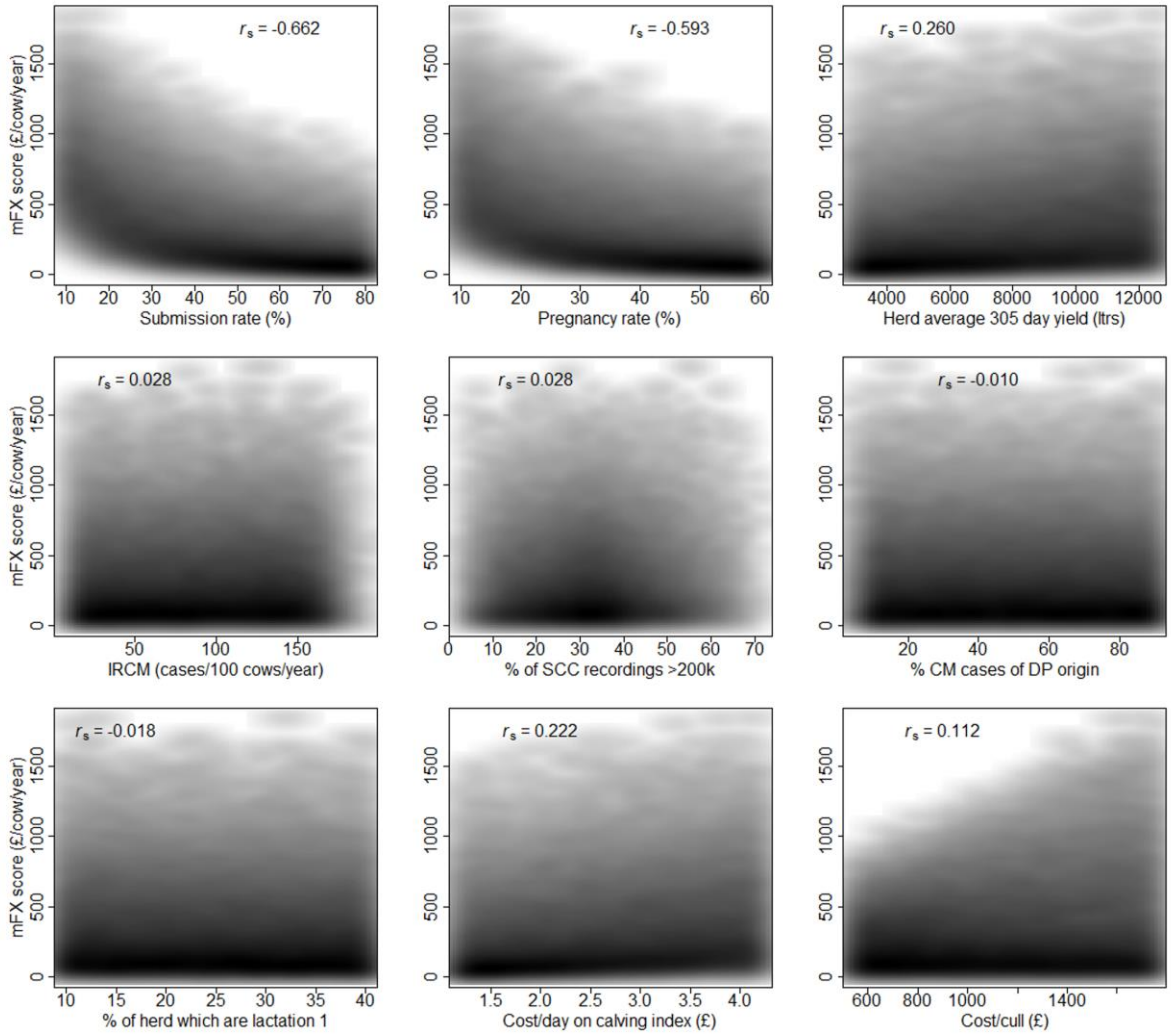


367

368

369 Fig. 2: Process for simulation of udder health history throughout a lactation. Solid black lines
370 indicate process flow, and dotted lines indicate that information from the source of the line is
371 used in the step of the process to which the line leads (denoted by a diamond). Fig. 2a shows
372 the proportion of clinical mastitis cases in the dataset from Hudson et al. (2012) by days in
373 milk, split into likely dry period versus lactation origin using data from Green et al. (2002).

374



375

376 Fig. 3: High-density scatterplots showing associations between overall fertility outcome and
377 herd-level input variables. Darker colours indicate higher densities of points. r_s, Spearman
378 rank correlation coefficient; FERTEX, modified FERTEX score (representing overall herd
379 fertility outcome); IRCM, incidence rate of clinical mastitis; SCC, Somatic cell count; CM,
380 clinical mastitis; DP, dry period.

381

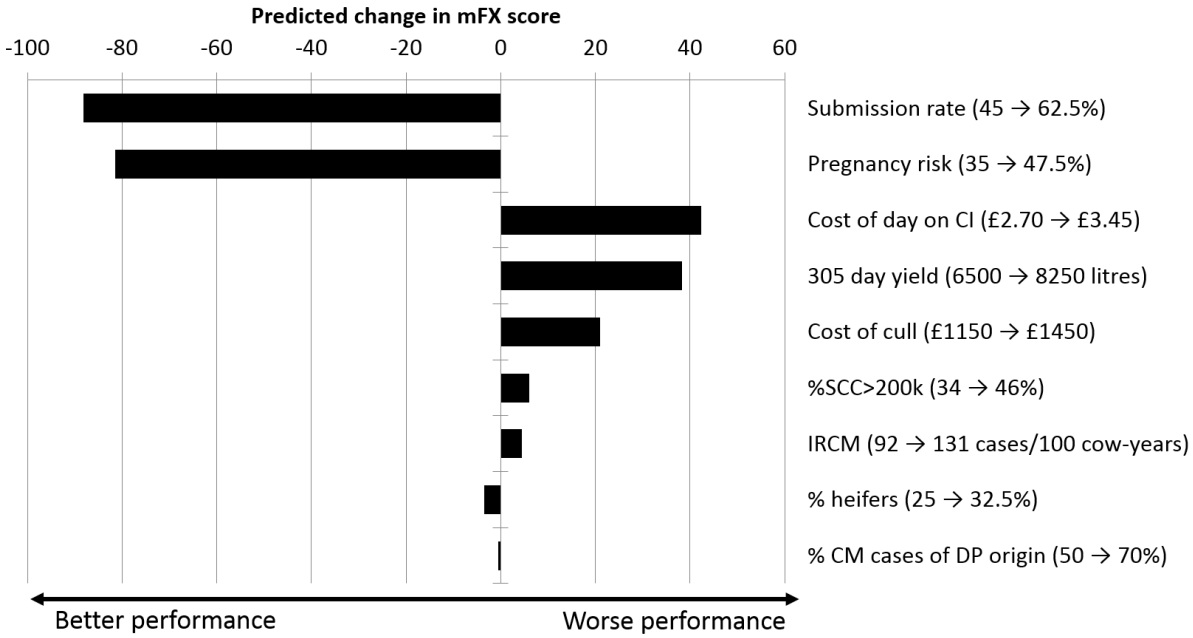
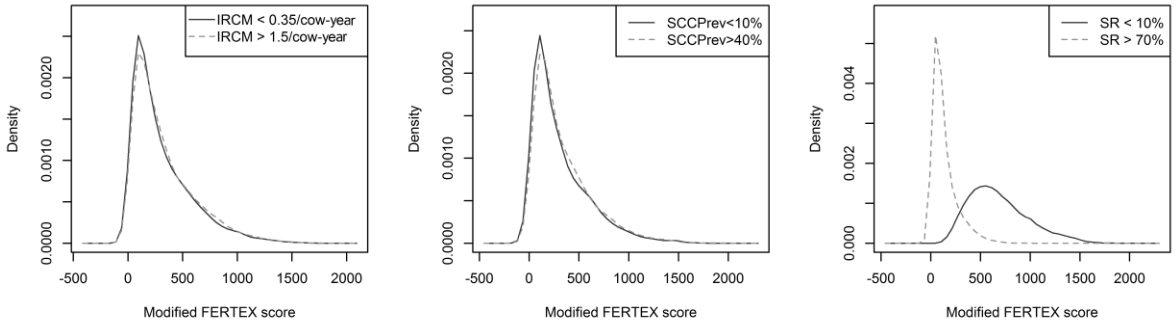


Fig. 4: Predicted effect of an equivalent increase in each input parameter on overall fertility. Tornado plots showing the predicted effect of increasing each input parameter in turn by a value representing 25% of the range of its input distribution from the median value, while the other input parameters are held at their population medians. The input parameters are listed on the right hand side of the graph, and the change in each input (from median to upper quartile) is given in parentheses. For example, the top bar shows that the predicted effect of moving from a submission rate of 45% (the median of the input distribution for this parameter) to 62.5% (the upper quartile of the input distribution) would be a decrease of just under £90/cow/year in herd mFX score.

Note: for the proportion of recordings where SCC>200k parameter (which was the only input not drawn directly from a uniform distribution), the change in the parameter (+12.4%) represented 25% of the 95% coverage interval of the distribution of this parameter.



396

397 Figure 5: Kernel density plots for simulated herds with extreme input parameter values.
398 Kernel density plots showing distribution of modified FERTEX score (as a measure of
399 overall fertility outcome) for herd with extreme values for: (a) IRCM (incidence of clinical
400 mastitis in cases/100 cows/year: IRCM<0.35 cases/cow-year, solid line; IRCM>1.5
401 cases/cow-year, dotted line); (b) proportion of somatic cell count recordings >200k
402 (SCCPrev; proportion <10%, solid line; proportion >40% dotted line); and (c) submission
403 rate (SR; submission rate <10%, solid line; submission rate >70%, dotted line)

404

Appendix A: Supplementary materials and methods

Discrete time survival model

The discrete time survival model on which the simulation model is based was described in Hudson et al. (2012), but is briefly summarised below:

The model was fitted using data from 80 dairy herds from across England and Wales. The main aim was to evaluate associations between reproductive performance and mammary gland health. A wide variety of potential explanatory variables relating to each cow's clinical mastitis (CM) and individual cow somatic cell count (ICSCC) history were used, along with other variables that potentially confound any relationship with reproduction (e.g. stage of lactation, 305d milk yield, lactation number, season etc.). A discrete time survival model was constructed within a multilevel framework, to account for correlations between lactations from the same cow and between cows in the same herd. A discrete time survival model is effectively a logistic regression model which predicts the probability that the event of interest (in this case, conception) occurs during each (discrete) unit of time (in this case, each 2-day period of a cow's lactation). The model took the conventional form:

$$\mathrm{Preg}_{tij} \sim \mathrm{Bernoulli}\left(\mathrm{mean} = \mu_{tij}\right)$$

$$\ln\left(\frac{\mu_{tij}}{1-\mu_{tij}}\right) = \alpha + \beta_1 \ln\mathrm{DIM}_{tij} + \beta_2 \left(\ln\mathrm{DIM}_{tij}\right)^2 + \boldsymbol{\beta}_3 \mathbf{X}_{tij} + \boldsymbol{\beta}_4 \mathbf{X}_{ij} + \boldsymbol{\beta}_5 \mathbf{X}_j + u_{ij} + v_j \quad (1)$$

$$v_j \sim \text{normal distribution } (0, \sigma_v^2) \quad (2)$$

$$u_{ij} \sim \text{normal distribution } (0, \sigma_u^2) \quad (3)$$

where t represents a 2-day risk period and i and j the i^{th} cow in the j^{th} herd; μ_{tij} the fitted probability of Preg_{tij} (the outcome of the i^{th} cow in the j^{th} herd becoming pregnant during risk period t); $\ln\mathrm{DIM}_{tij}$ the natural logarithm of days in milk at the beginning of risk period t; α the regression intercept; β_1 and β_2 the coefficients for the terms representing days in milk; $\mathbf{X}_{tij}$ the vector of risk period level covariates and $\boldsymbol{\beta}_3$ the corresponding vector of coefficients for

covariates $\mathbf{X}_{tij}$; $\mathbf{X}_{ij}$ the vector of cow-level covariates and $\boldsymbol{\beta}_4$ the corresponding vector of covariates of coefficients $\mathbf{X}_{ij}$; $\mathbf{X}_j$ the vector of herd-level covariates and $\boldsymbol{\beta}_5$ the corresponding vector of coefficients of covariates $\mathbf{X}_j$; u_{ij} the random effect to reflect variation between individual cows and v_j the random effect representing variation between herds, with σ_u^2 and σ_v^2 the variances of the normal distributions of the respective random effects terms.

Explanatory variables from this model which were significantly associated with the probability of a cow becoming pregnant during a 2-day risk period were used as input parameters for the simulation in this study, with the exception of year of calving (as this effect was not considered relevant) and three ICSCC related variables which had very small associations with the outcome (which were omitted for model parsimony). Readers are referred to the original publication (Hudson et al., 2012) for estimated model coefficients and interpretation.

Correlations between input parameters

The possibility that correlations between input parameters would affect the simulation outcome was investigated using the following method. Distributions of these input parameters for each of the 80 herds in the original dataset from Hudson et al. (2012) were evaluated. Assessment of the univariate distribution of each parameter in turn showed that the ranges of the parameters across herds were very similar to those chosen for the uniform input distributions shown in Table 1, and that many of the inputs did not appear normally distributed. As it was plausible that all inputs were jointly correlated in a complex fashion (and clear that few approximated a normal distribution), attempting to fit a parametric multivariate distribution to the data was considered inappropriate. Instead, a non-parametric approach was taken, whereby the simulation exercise was repeated using the observed joint

distribution of the parameters across the herds was used as simulation inputs, so that at each iteration of the simulation the set of observed input parameters for one of the 80 herds was used as the input for the simulation model. This process was also repeated using the joint distributions of input parameters observed for each herd-year (i.e. for each herd in each year) in the original dataset (n=435).

Repeating the simulation and analysis using the observed joint input distributions from the original dataset (instead of those described in Table 1) affected the results of the univariate analyses, but multivariate regression analyses produced similar results to those generated using independent uniform input distributions. Although the regression coefficients for both udder health related input parameters increased slightly (and the predicted effect of IRCM became the larger of the two), the predicted effect of changes in these parameters remained much smaller than the predicted effects of changes to the key drivers of mFX score. Supplementary Figure 1 shows the tornado plot generated using the observed joint input distributions of herd-years from the dataset; the joint distribution at herd level produced an almost identical plot. It therefore appears that the choice between these alternative input distributions would not have a substantial impact on the biological interpretation of the results of this study, and the results reported in the main manuscript were derived from the original uniform input distributions.

Generation of clinical mastitis and individual cow somatic cell count history for a simulated lactation

For CM, the herd-level input parameters were the incidence rate of CM and the proportion of CM cases resulting from intramammary infection during the dry period. In order to use these parameters to predict occurrence of CM as a binary event for each two-day risk period, a value for the number of DIM at each case of CM was extracted from the 80-

herd dataset: this determined the distribution of cases of CM over the course of lactation. A total of 67,994 cases of CM were included in this analysis. Data from Green et al. (2002) were then used to attribute the proportion of cases at each two-day period through lactation as either dry period or lactation origin, with a very high proportion of cases in early lactation being attributed to the dry period (Figure 2a), and a very high proportion of cases in late lactation attributed as lactation origin. These results were then used to calculate the proportion of all dry period origin cases and of all lactation origin cases which occurred at each two-day risk period. For each herd simulated, the input parameters were used to determine the separate incidence rates for dry period and lactation origin CM (by multiplying the overall incidence rate by the proportion of cases of dry period origin). This allowed prediction of the probability of the occurrence of either dry period origin or lactation origin CM at each two-day risk period during the lactation: the simulation model then assigned events by drawing from a binomial distribution based on the calculated probability of CM at each risk period.

In order to simulate ICSCC history, it was assumed that the cow would have a first milk test day of the lactation at a random stage within the first 30 DIM (so that DIM at first test day was drawn from a uniform distribution between 0 and 30), and would have test days at regular 30 day intervals after this. ICSCC was treated as a binary variable, such that the cow could occupy one of two states; infected (ICSCC>200k) or uninfected (ICSCC<200k). The herd-level input parameters were then used to determine the cow's status at the first recording of lactation (a draw from a binomial distribution with probability equal to the overall proportion of cows with a first ICSCC of lactation >200k), and the likelihood that her status will change at each subsequent test day.

Combining reproductive outcomes to a single lactation-level measure

To simplify analysis of the results of the simulation, a single outcome representing herd fertility performance was required. For each simulated herd, the proportion of the herd which reached 300 DIM without becoming pregnant was calculated (this was used as a proxy for the rate of fertility-associated culling) along with the mean number of DIM at conception (which was converted to a mean herd calving index by adding 282 days for gestation). These were then combined by comparing each to a selected baseline value (345 days for calving index and 0% for 300 day failure to conceive rate), applying a cost per unit deviation from the target (with unit cost for each represented as herd-level input parameters) and summing the total cost per cow to create a modified 'FERTEX' (mFX) score for each herd (Esslemont and Kossaibati, 2002). The baseline values for calving index and failure to conceive at 300 DIM were intentionally set at very low levels to avoid herds which performed better than the baseline level (and therefore had negative mFX scores). Although this mFX score represented an appropriate single outcome measure for this study, the absolute value of mFX score for each simulated herd would therefore not reflect true recoverable loss due to infertility (although changes in mFX score would be realistic).