Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Bayesian Logistic Disease Progress Curves for Modeling Plant Disease Epidemics

Claudio Dias da Silva
Kansas State University
Aidan Kerns
Kansas State University
Erick DeWolf
Kansas State University
Trevor Hefley
Kansas State University
Kelsey Andersen Onofre
andersenk@ksu.edu

Kansas State University

Method Article

Keywords: Plant disease epidemiology, Bayesian inference, Uncertainty quantification, disease progress curves

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11484962/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Bayesian Logistic Disease Progress Curves for Modeling Plant Disease Epidemics

da Silva Jr., C. D.[1,2], Kerns, A[2], De Wolf, E.[1], Hefley, T. J.[2], Andersen Onofre, K.[1*]

Affiliations

[1]Department of Plant pathology, Kansas State University, Manhattan, KS, 66506

[2]Department of Statistics, Kansas State University, Manhattan, KS, 66506

*Correspondence: andersenk@ksu.edu; cdiasda@ksu.edu; thefley@ksu.edu

Abstract

Background: Logistic disease progress curves have been classically used to characterize plant disease epidemics over time. When observations are few and information is limited, joint estimation of their three parameters (all three as unknown quantities) including maximum disease potential (K), apparent rate of growth (r), and initial disease level (y_0) can be difficult. The standard approach in plant pathology is to fix the upper asymptote at a pre-specified value (usually 1) which simplifies estimation while excluding uncertainty about maximum disease potential and potentially underestimating r. We explain a Bayesian logistic disease progress curve framework that combines biologically informed priors with disease severity observations to estimate all three parameters, while appropriately accounting for uncertainty in predictions and derived quantities (i.e., functions of model parameters).

Results: We applied the model to two case studies involving wheat stripe rust (*Puccinia striiformis f. sp. tritici*), each with different levels of information about epidemic development. Posterior contrasts with probabilistic inference provided a comparative epidemic framework for model parameters and derived quantities. In case study 1, posterior contrasts provided strong evidence of difference in r between treatments, whereas evidence for difference in K was weaker when observations near the asymptote were limited. For case study 2, where there was overall lower disease severity, the posterior distribution of K closely resembled its prior, whereas posterior distributions of r and y_0 shifted toward lower values. These results illustrate how the information available related to observed disease severity over time affected the relative contributions of prior information and data to parameter inference. In-sample coverage of 95% posterior predictive intervals was 96.7% and 100% for case studies 1 and 2, respectively. Coverage was 44.4% when

predictions were evaluated against observations from another season, indicating limited predictive transfer.

Conclusion: We propose and demonstrate a biologically informed Bayesian framework that supports characterization and probabilistic comparison of incompletely observed plant disease epidemics and calculation of derived quantities, such as area under disease progress curve, while explicitly propagating uncertainty. Biologically informed prior distributions supported joint estimation of all three parameters when observations are limited. Publicly available code facilitates application by the plant science community.

Keywords: Plant disease epidemiology; Bayesian inference; Uncertainty quantification; disease progress curves

Background

Nonlinear mathematical models are valuable tools for studying dynamic biological processes in plant science because of their flexibility to describe complex patterns and because of their biologically interpretable parameters [1]. Nonlinear models have been widely used in plant pathology for the study of plant disease epidemics. These include disease progress curves (DPCs), which are commonly used to study plant pathogen population dynamics and disease progression over time [2]. Several DPC models have been developed over the years with varying levels of utility and diverse applications [3]. Such models have played a pivotal role in studying the influence of management strategies on plant disease epidemic development [1, 4].

DPCs have foundational biological meaning, and their parameters represent different steps of plant disease epidemic development. Logistic DPCs (LDPCs), for example, have been used to study disease epidemics, where continuous development within a crop cycle (e.g., multiple cycles of pathogen reproduction) results in a sigmoidal shaped curve [4, 5]. In this model, there are three unknown quantities (parameters), the apparent rate of growth (r), the initial disease level (y_0), and the maximum disease potential (K). Over the years, these parameters have been associated with the effect of different management strategies. Chemical control, for example, has been shown to affect both initial disease level and apparent rate of growth [3, 6].

Although DPC parameters can help characterize the effects of management practices on overall disease development, estimating them is challenging when datasets are small or provide limited information about the disease development trajectory. In fact, the small dataset problem has been well documented in the plant sciences [7-9]. Such problems often lead to practical identifiability issues, where it is difficult to determine the parameter value from the available data, resulting in high levels of uncertainty and correlation among parameters [10-12]. To deal with

such issues, strategies such as modifying the experimental design and/or data collection, or model simplification, are available [13].

However, such strategies might not always be feasible or appropriate and represent only a subset of the tools that could be used to remedy the problem. Adjusting experimental design and/or data collection inevitably involves collecting more data, which is not always possible, and does not produce an immediate solution for the available data. Model simplification can be applied to the available data, producing an immediate solution. In plant disease epidemiology, such strategies have been widely used to fit DPCs, with common practices involving linearization through transformation and fixing parameter values instead of estimating then [2, 4, 14].

For LDPCs, model simplification often involves the assumption that the maximum disease potential is 100% ($K = 1$). Although it effectively helps with practical identifiability issues, such an assumption imposes a very strong constraint (the strongest possible) on K, and as a consequence, it can lead to underestimation of r, which is often of primary interest when comparing disease management strategies [15]. In fact, it has been shown that this assumption often does not hold across environments, and that K is highly variable across different disease systems [16, 17]. Another important caveat is the loss of inferential capability on K, with little being known regarding how different management strategies affect this parameter.

A third strategy to work with practical identifiability issues that occur when data is informationally deficient is the use of additional information and constraints imposed by prior distributions in Bayesian inference. In Bayesian inference, prior distributions are combined with observed data yielding posterior distributions that reflect both sources of information [18]. Informative priors derived from external sources of information (literature, external data, and/or

expert knowledge) can constrain the parameter space to plausible regions, in other worlds, supply additional information where available data is deficient [19].

In this paper we present a Bayesian LDPC model with the goal of jointly estimating (i.e., treating all as unknown quantities) all three model parameters (y_0, r, and K) for a disease dataset with different levels of information deficiency. We used data from field trials designed to study different management strategies for wheat stripe rust, in which disease severity (percentage of leaf area diseased) was evaluated over time. Additionally, we demonstrate the calculation of derived quantities (i.e., functions of model parameters) commonly used in plant pathology such as area under the disease progress curve (AUDPC) and contrasts to compare posterior distributions of parameters and derived quantities, respectively. We present Bayesian inference as a method for plant pathologists and plant scientists working with nonlinear models under limited data or incomplete observation scenarios. These methodologies are presented in associated publicly available, annotated code for use by the research community.

Methods

Case study experimental design

Two field experiments were conducted during the 2022/2023 [Season 1] and 2023/2024 [Season 2] growing seasons to test different management strategies for wheat stripe rust. Both were located in the Kansas River Valley Experimental Field located in Rossville, KS (39.0769, -95.7689) and received supplemental irrigation. In both seasons, experiments were configured in a split-plot design with three replications, with fungicide as the main factor (12 levels; Table 1) and wheat cultivar the second (KS-180B, SY Monument). Fungicides differed by their active ingredient, with five different formulations, and application timings, being at Feekes growth stage

7/Zadoks 36 (FGS7) and Feekes growth stage 9/Zadoks 39 (Table 1). Cultivars differed by their reaction to stripe rust, with KS-180B presenting a susceptible reaction, and SY Monument a moderately resistant reaction.

Disease evaluation

Wheat stripe rust incidence and severity were assessed at regular intervals following flag leaf applications through harvest. Assessments were conducted at 11-day intervals in Season 1 and 7-day intervals (weekly) in Season 2. Disease incidence was recorded as the percentage of diseased plants per subplot, and disease severity was recorded as the percentage of diseased leaf area on the upper canopy of affected plants (flag leaf and flag leaf minus one, the top two leaves of the plants).

Bayesian Logistic Disease Progress Curve

Model specification

A Bayesian Logistic Disease Progress Curve (Bayesian LDPC) model was constructed to model the progress of disease severity over time. Disease severity, measured as a percentage, was converted into proportion for model fit ($y = \frac{sev}{100}$). The data was assumed to follow a beta distribution whose support naturally fits proportions (bounded between 0 and 1). Since this distribution does not handle zero's and one's, the rescaling function $y^* = \left[y * (N-1) + \frac{1}{2}\right]/N$ (where N represents the sample size) was used to shrink the interval from [0, 1] (including 0 and 1) to (0, 1) (between 0 and 1, excluding the boundaries) [20]. The statistical model can be written out as:

$$y_{ij}^{*} \sim beta(\mu_{ij}, \phi_i), \quad (1)$$

$$\mu_{ij} = \frac{K_i}{1+\left(\frac{K_i - y_{0i}}{y_{0i}}\right) * e^{-r_i * t_j}}, \quad (2)$$

$$y_{0i} \sim uniform(0.001, 0.1), \quad (3)$$

$$K_i \sim beta(6, 2), \quad (4)$$

$$r_i \sim uniform(0, 0.5), \quad (5)$$

$$\phi_i \sim gamma(24, 2). \quad (6)$$

164

165 The model was defined for each *i*th fungicide treatment level, with the epidemic developing
166 over the *j*th days after the application made at FGS7. Eq. 1 describes the data model, where a beta
167 regression parameterization of the beta distribution, with expectation μ_{ij} and dispersion ϕ_i, was
168 adopted [21]. Eq. 2 describes the process model behind the data model expectation (μ_{ij}), where
169 the mathematical equation of the LDPC was adopted with three parameters, the maximum disease
170 potential (K_i), the apparent rate of growth (r_i), and the initial disease level (y_{0i}), all of them
171 considered treatment specific. Eq. 3 to 6 describe the prior distributions for each model parameter.
172 Prior distributions were specified using fundamental biological knowledge behind the LDPC
173 parameters and information available in the literature. In the following paragraphs, we explain in
174 detail the rationale behind each curve parameter prior.

175 The maximum disease potential (K) is a parameter representing the curve asymptote, the
176 final state of the epidemic curve, which y approaches as time passes. In some cases, the maximum
177 disease intensity has been considered as 100% ($K = 1$) for simplification [6]. However, it has been
178 noted that severity of leaf spots and rust diseases rarely achieve such levels [22]. In fact, across
179 different pathosystems, K is highly variable. Examples of such variability include the modeling of

incidence of tobacco viruses, where K ranged from 0.01 to 1 [23], and the modeling of plant mortality caused by *Phytophthora cinnamomi*, where K ranged from 0.52 to 1 [24]. For stripe rust, estimates of maximum disease severity have been shown to range from ~ 0.2 to ~ 0.7 [25]. We started from the idea that K could only assume values between 0 and 1, considering it represents a maximum proportion. A natural distribution for modeling proportions, the beta distribution, was adopted as prior. Although K could clearly assume small values (e.g., close to 0.2), based on field observations [data not shown], it more often assumes values close to 0.7 for this pathosystem. We then used this information on maximum disease severity to derive values for the beta shape parameters that would yield a beta distribution with mean of 0.75 and a variance of 0.02. Such mean and variance produce a beta prior whose 95% quantiles are 0.42 and 0.96 (Eq. 4, Supplementary Figure 1).

The apparent rate of growth (r) is a parameter representing how fast the disease epidemic increases over time, from the initial state (initial disease level, y_0) toward the asymptote (maximum disease potential, K). Examples of estimated r values for wheat stripe rust epidemics in literature range from ~ 0.03 to ~ 0.2 [26-28], however, such values have been obtained under the assumption of $K = 1$. A natural lower bound for this parameter is 0, because a negative value would lead to disease decrease over time. When r is 0, the LDPC only propagates y_0 over the time of the epidemic. For the upper bound, intuitively, 0.2 would be a natural choice, however, the assumption of $K = 1$ could lead to underestimation of r [15], which lead us to assume an upper bound of 0.5, giving some freedom for higher growth rate values. For r, we opted for the uniform distribution as prior, since it constrains the parameter within the desired biologically relevant boundaries, without attributing higher densities to any specific region, that is, the prior density is constant over the interval 0 to 0.5, and values of $r > 0.5$ or $r < 0$ cannot occur (Eq. 5, Supplementary figure 1).

The initial disease level (y_0) is a parameter representing the initial state of the epidemic, or the disease intensity at time 0. It is the counterpart of K, and similarly, it has been often assigned a fixed value, being either the initial measure or a very small number. Considering that y_0 is usually represented by very small values, close to 0, a semi-informative prior was used, in which bounds were defined to constrain the parameter value, but considerable room for variation was allowed. A uniform distribution was defined for y_0, with boundaries targeting a very low initial level (0.1%, 0.001), and a considerably higher initial level (10%, 0.1) of disease severity, and values $y_0 > 0.1$ and $y_0 < 0.001$ not being possible to occur (Eq. 3, Supplementary figure 1). These priors also make the violation of $y_0 \leq K$ relationship unlikely, ensuring the curve will not decrease.

Model implementation

The model was implemented using the statistical software R with the package rjags (JAGS – Just Another Gibbs Sampler) for posterior sampling [29, 30]. The MCMC (Markov chain Monte Carlo) algorithm was conducted with 3 chains, each chain had 15000 iterations with a burn-in of 5000. MCMC convergence and sampling efficiency were assessed through diagnostics such as visual inspection of trace plots, the potential scale reduction factor ($\hat{R}$), and effective sample size (ESS).

Statistical analysis: Model checking and Inference

With the Bayesian LDPC fully specified, prior predictive checks were performed to assess whether the specified priors and observation model generated biologically plausible disease severity patterns and disease progress trajectory. These checks involved visual inspection of the prior predictive mean and 95% prior predictive interval for simulated observations (y^*_{prior}),

together with 50 expected disease development trajectories generated from randomly selected prior parameter draws (μ_{prior}).

Following prior predictive checks, the model was fit to data and posterior distributions were used to characterize and compare epidemics, and quantify uncertainty in predictions and derived quantities. The analysis was organized into two case studies. For both case studies, data from Season 2 were utilized. In case study one [CS1], data from two epidemics from the treatments with Prosaro applied at FGS 7 and FGS 9, for the susceptible cultivar (KS-180B) were utilized. Observations for application at FGS 7 produced a sigmoid disease development pattern, while observations for application at FGS 9 lacked information, especially at the curve asymptote. In case study two [CS2], data from only one epidemic, the healthy check treatment for the moderately resistant cultivar (SY Monument) were utilized (Table 1; Supplementary Figure 2). CS2 was designed to demonstrate model fit when information on the curve shape was absent in addition to a limited number of observations. LDPC model parameters were jointly estimated and posterior predictive curves were obtained for both case studies. Additionally, for CS1, two derived quantities, the area under the disease progress curve (AUDPC, Table 2) and the maximum instantaneous rate of growth (r_*; Table 2), were obtained to summarize and compare the disease progress curves. r_* represents the highest slope of the sigmoid curve before the rate of disease progress begins to decrease. Posterior distributions of parameters and derived quantities were compared through posterior contrasts (Table 2). The probability that the contrast was bigger than zero ($P(a > b) = P(\Delta > 0)$) was obtained for probabilistic inference. A probability of 0.5 ($P(\Delta > 0) \cong 0.5$) indicates evidence of equal posterior support in both directions, a probability greater than 0.5 ($P(\Delta > 0) > 0.5$) is evidence of $a > b$, and a probability less than 0.5 ($P(\Delta > 0) < 0.5$) is evidence of $a < b$, with the strength of evidence increasing as probability

values approach 0 or 1. Here, a and b denote quantities being compared, either parameters or functions of parameters.

Posterior predictive checks were performed by assessing the calibration of the 95% posterior predictive intervals for the data used for model fitting (in-sample posterior predictive check). For CS1, predictive performance was further evaluated out-of-sample (i.e., data not used for model fit) using data from Season 1 for the same treatment and variety [31].

Results

MCMC diagnostics

For both case studies, trace plots showed overlapping chains fluctuating around stable levels, with no sustained upward or downward drift in parameter values across iterations (Supplementary Figure 3 and 4). The potential scale reduction factor ($\hat{R}$) ranged from 1.00 to 1.03 in CS1 and 1 for all parameters in CS2. Minimum effective sample sizes were 769 and 3108 for CS1 and CS2 respectively, across the monitored parameters (Supplementary table 1). Together, these diagnostics supported satisfactory chain mixing and sampling efficiency.

Prior predictive checks

For simulated observations (y^*_{prior}), a prior predicted curve was obtained for a period of 55 days without conditioning on observed data (Figure 1a). The prior predicted mean severity curve (proportion) presented a short initial lag-phase, and rapid growth towards the asymptote, with growth slowing around 30 days. The obtained 95% prior predictive interval was narrower at early stages of the prior predicted curve, followed by substantial widening at later stages as days progressed from 1 to 55, reflecting the combined uncertainty from parameter priors and

observational model. The 50 sampled expected disease development trajectories (μ_{prior}) were consistent with biological expectation. Some trajectories were flat, representing little disease increase over the prediction period, while others quickly increased, approaching the asymptote (Figure 1b). Together, these trajectories and the simulated disease severity patterns supported the biological plausibility of the specified model and priors.

Case study 1

Inference on model parameters

Inference on posterior distribution for model parameters was made on the maximum disease potential (K) and the apparent rate of growth (r). Upon visual inspection of posterior distributions using histograms, evidence of difference on the parameter r between application timings was observed, once they have almost no overlap (Figure 2). For applications made at FGS 7, the expected value of the posterior distribution (thereafter referred as posterior mean) was 0.21, with 95% credible interval (CrI) from 0.17 to 0.23. For applications made at FGS 9, the posterior mean was 0.08, with 95% CrI from 0.017 to 0.13. All posterior samples of the contrast were positive, yielding a probability of the contrast being above zero of 1 ($P(\Delta > 0) \approx 1$; Figure 3; Table 3), providing strong evidence that the rate of disease progress was faster for applications made at FGS 7.

Evidence of difference was also observed for K, although not as strong, with posterior distributions presenting overlap (Figure 2). For applications made at FGS 7, the posterior mean was 0.89, with CrI from 0.74 to 0.98. For applications made at FGS 9, the posterior mean was 0.75, with 95% CrI from 0.44 to 0.96, being broader and indicating higher uncertainty. Posterior samples of the contrast were mostly above zero (Figure 3), the probability that the contrast is above

zero was of 0.81 ($P(\Delta > 0) = 0.81$), providing evidence that K for applications made at FGS 7 was greater than K for applications made at FGS 9 besides considerable uncertainty (Table 3).

Posterior predicted curves

Posterior predicted curve and 95% posterior predictive interval (PPI) were obtained for both application timings. Predictions were made over 55 days. The posterior predicted curve for the application made at FGS 7 quickly increased and approached the plateau (stationary phase) within this time period, while the curve for the application made at FGS 9 starts its exponential-phase at the end of the time period. The curve for FGS 9 has higher uncertainty, particularly beyond the observation period, which, in part, reflects the uncertainty propagation from K. Although 95% PPI overlap was observed initially, the curves begin to diverge after 20 days, when overlap decreases (Figure 4).

For the in-sample posterior predictive check, the 95% PPI contained 96.7% of the observed disease severity overall, with coverage of 100% and 93.3% for FGS 7 and FGS 9 respectively. For out-of-sample posterior predictive check using observations from Season 1, the 95% PPI contained 44.4% of the observations overall, with coverage of 22.2% and 66.7% for FGS 7 and FGS 9 respectively. Coverage was substantially below the nominal 95% level in a different season, particularly for FGS 7, indicating limited predictive transfer capability across seasons.

Derived quantities

The posterior distributions for AUDPC were calculated for each posterior draw of the mean disease progress curve from day 1 to day 55, using the equation in Table 2. For applications made at FGS 7, the posterior mean of the AUDPC was 21.72, with 95% CrI from 18.83 to 24.09. For applications made at FGS 9, the AUDPC posterior mean was 5.21, with 95% CrI from 2.08 to 9.40. Posterior AUDPC distributions had no overlap (Figure 5), reinforced by a probability that

the contrast was above zero of 1 ($P(\Delta > 0) \approx 1$; Table 3), providing strong evidence that the posterior AUDPC for FGS 7 was greater than the posterior AUDPC for FGS 9.

The posterior distributions for r_* were calculated for each posterior draw from K and r for each treatment. Applications made at FGS 7 had a posterior mean of 0.04, with 95% CrI from 0.04 to 0.05. For applications made at FGS 9, the posterior mean was 0.01, with 95% CrI from 0.003 to 0.03. Little overlap was observed between posterior distributions (Figure 6) for the two different application timings, with a probability that the contrast was above zero of 0.99 ($P(\Delta > 0) = 0.99$; Table 3). This was strong evidence that at the point of maximum rate of disease progress, the curve for applications made at FGS 7 was faster than the curve for applications made at FGS 9, agreeing with patterns observed in the posterior predicted curves (Figure 4).

Case study 2

Inference on model parameters

The disease severity data for CS2 remained near zero throughout the observation period, providing limited information about disease development (Supplementary Figure 1). This data was generated from treatment involving a moderately resistant cultivar and frequent fungicide application (Healthy check; Table 1). In such a scenario, little information regarding K was available in the observed data, which in turn lead to a posterior sample from this parameter essentially driven by the prior (Figure 7; Supplementary figure 1). In fact, K remained mostly unchanged, while r and y_0 shifted toward smaller values and became more concentrated compared to their priors (Table 4). The posterior mean of K was 0.75, with 95% CrI from 0.42 to 0.96, the same as the specified prior. Most of the information in the observed data was related to y_0, as the disease severity remained low over time for this treatment. The posterior mean of y_0 was 0.013,

with 95% CrI from 0.004 to 0.026, with small uncertainty (Figure 7). In this case study, the posterior distribution of r concentrated near zero, consistent with limited disease progression during the observation period. The posterior mean of r was 0.016, with 95% CrI from 0.0005 to 0.0509. The posterior distribution of r was broad, with a skew toward higher rates of disease progress (right-skewed; Figure 7).

Posterior predicted curve

Posterior predicted curve was obtained for a time period of 55 days. The wide 95% PPI represents high predictive uncertainty, that increases near the end of the time period (Figure 8). At day 1, the posterior predicted mean severity (proportion) was 0.0138, with 95% PPI from 10^{-14} to 0.1098. At day 55, the posterior predicted mean severity (proportion) was 0.0333, with 95% PPI from $2x10^{-7}$ to 0.1867. The posterior predicted mean increased 0.0195 in this time interval, while the 95% PPI widened from 0.1098 at day 1 to 0.1867 at day 55. For in-sample posterior predictive check, the 95% PPI contained 100% of the observed disease severity for the healthy check treatment.

Discussion

This paper demonstrates the use of Bayesian inference to characterize plant disease epidemics when observed data provide limited information about epidemic progress. We jointly estimated all three parameters (K, r, and y_0) from a logistic disease progress curve (LDPC) by combining prior information and observed data, while also propagating uncertainty to predictions and derived quantities. We used two case studies as examples with different levels of information available about epidemic development. In CS1, we demonstrated the method for data with two treatments, were one had limited information around the curve asymptote (FGS 9). In CS2, a more

363 drastic scenario, we demonstrated the method for data with low disease severity, with little
364 observed progression as a result of the applied treatment. In both case studies, posteriors are a
365 result of the combination of prior and observed information, however, the relative contribution of
366 these sources of information to parameter posteriors differed across cases. This difference became
367 clear when comparing posterior distributions of K obtained in CS1 and CS2, where for the latter,
368 prior and posterior distributions of this parameter were similar.

369 These findings highlight how disease development observed under different management
370 strategies affect the information available for individual LDPC parameters, especially for time-
371 limited epidemics, where K seems to be the most influenced parameter. Jointly estimating K, r,
372 and y_0 makes uncertainty about the asymptote explicit and propagates it to predictions and derived
373 quantities, including when inference on K relies predominantly on prior information. We adopted
374 the LDPC model, which conceptually matches the current understanding of the biological process
375 influencing the development of stripe rust within a wheat field [4, 5]. It should be noted that
376 biological knowledge regarding the pathosystem was the driver of model selection and
377 specification. Distributional assumptions for both prior and likelihood considered the nature of the
378 data being collected and pre-existing knowledge regarding the pathosystem. The selection of the
379 mathematical model considered the pathogen biology rather than the pattern observed in the
380 collected data [32].

381 For CS1, observations were collected from the susceptible cultivar, with treatments
382 consisting of two different application timings (FGS 7 and FGS 9) of the same fungicide product.
383 In this case study, application timing affected both K and r, but with different intensities. For r,
384 there was strong evidence that the epidemic for applications made at FGS 7 progressed much faster
385 than the epidemic when applications were made at FGS 9. These results are consistent with

findings that chemical control reduces the rate of disease progress [3, 4, 6]. In this specific pathosystem, applications at FGS 9 are often recommended to manage the disease [33, 34]. For K, there was evidence that applications made at FGS 7 allowed for a greater maximum disease potential than FGS 9, likely due to the loss of residual efficacy of the fungicide product applied at the earlier growth stage [35]. Although inference on this parameter is not widely studied, it has been shown that chemical treatments often lead to a smaller K [36, 37], and that K can also be affected by the cultivar reaction to disease [24]. Broader posterior distributions of r and K for FGS 9, were consistent with the limited information available about disease development, particularly for the upper asymptote (K). Parameter uncertainty propagated into the derived quantities and posterior predictions for this treatment (broader posteriors/intervals). For posterior predictions, although informative priors can constrain parameter uncertainty, posterior predictive intervals also incorporate observation-level variability, contributing to wider PPI [38].

For CS2, observations were collected from the moderately resistant cultivar, and treatments consisted of several fungicide applications to maintain low disease severity. These data were chosen to show the effect of very limited information about K, consistent with practical non-identifiability [10-12, 15, 39], on the joint estimation of the LDPC parameters using Bayesian inference. While the posterior distributions for y_0 and r were influenced by the information from the observations, shifting from their priors, the posterior distribution of K closely resembled its prior, indicating that the observed disease trajectory contributed little additional information. Bayesian inference enabled joint parameter estimation by incorporating biological knowledge through the prior distributions, however, inference on K remained predominantly dependent on that information. Obtaining a posterior distribution, therefore, did not imply that K had been identified from the observed data [19]. This result highlights the importance of biologically

409 justified priors and of assessing the sensitivity of inference to their specification, particularly for
410 parameters weakly informed by the observed data.

411 Prior distributions for this study were built on a combination of information gathered from
412 the literature as well as from field observations from studies not included in this analysis. When
413 specifying priors, the most important consideration is that they should not contain the same
414 information as the likelihood (i.e., the data) [40, 41]. Specifying prior distributions makes
415 parameter estimation less sensitive to the information in the data, which may be limited [40]. For
416 K, r, and, y_0 we started with a probability distribution based on the intrinsic properties of the data
417 (i.e., proportions, bounded between 0 and 1), and complemented this with additional information
418 based on available field trial data on wheat stripe rust. There is abundant information about LDPC
419 models available in literature across many pathosystems, which can be used to specify weekly
420 informative priors [2-6, 14-17]. For K, fixing the parameter at 1 (a very common practice in plant
421 pathology) would not allow information from the collected data to contribute to estimation. Even
422 a weakly informative prior such as uniform(0, 1), would provide valuable additional information
423 by constraining the parameter space (i.e., values the parameter can assume) into a biologically
424 realistic range. In this study, we demonstrated that the estimation of K is advantageous for a
425 researcher using the LDPC model. Estimation of K allowed inference about this parameter and
426 reduced the risk of underestimating r, as would happen by assuming $K = 1$ in most cases [15], as
427 CS1 and CS2, for example. Additionally, by estimating this parameter, uncertainty was properly
428 accounted for and propagated into derived quantities and posterior predictions.

429 Bayesian models allow the implications of specified priors and model assumptions to be
430 examined through prior predictive checks [42]. These checks involve simulating observations from
431 the specified model before conditioning on observed data and assessing whether the simulated

patterns are biologically plausible [31]. In some cases, prior predictive distributions can also be used for forecasting when local data is unavailable and relevant information is available from other sources [43]. In the present study, simulated disease severity and expected disease development trajectories ranged from slow disease increase to rapid epidemic development, consistent with our biological expectations. These checks supported the biological plausibility of the specified model and priors. Prior predictive checking therefore provides an opportunity to identify implausible model implications before fitting the model to new observations [31].

Additionally, we also conducted posterior predictive model checks by testing the calibration of the 95% PPI. Results for this check must be carefully considered based on the number of observations available. Treatments in Season 2 had 15 total observations, while in Season 1 there were 9, with 3 replications (blocks) at each sampling date, which means that within each season, an observation carries 6.7% and 11.1% of the total, respectively. For out-of-sample validation, for example, one observation not covered by the 95% PPI, would result in a coverage of 88.9% in the calibration check. For CS1, where out-of-sample calibration was conducted, a very poor transfer capability was identified. However, the effect of season was never accounted for within the model, which might explain such results. This example demonstrates that although our prior distributions are general assumptions for our pathosystem, the posterior distributions are a result of both prior and data (likelihood) information, and therefore, will not generalize for all scenarios, if they are not accounted for in the model. For generalization, a Bayesian model per se has the advantage of using prior information for estimation, but still requires proper model specification, even if broad general priors were specified.

Derived quantities were calculated as a demonstration of possible ways to summarize and compare disease progress curves. We chose to calculate the absolute/raw area under disease

455 progress curve (AUDPC), because it is one of the most widely used derived quantities in plant
456 pathology, with straightforward computing implementation [3, 44, 45]. Other options of derived
457 quantities such as K_* and r_* have also been used to summarize and compare different disease
458 progress curves [37]. These two derived quantities have similar units as the parameters involved
459 in their calculation, and their dimension is not dependent on the duration of the epidemic (i.e.,
460 period of observation), which makes comparison across studies straightforward. In a Bayesian
461 model, the derived quantity is calculated for each draw of the posterior distribution of parameters
462 involved in its calculation, resulting in a posterior distribution of its own, and ensuring that
463 uncertainty is properly propagated. In CS1, for example, the AUDPC was calculated based on the
464 analytical solution of the integration (Table 2), using the posterior distributions of K, r, and y_0,
465 therefore, propagating the uncertainty from μ into the AUDPC posterior distributions.

466 In both case studies, block effects were omitted to simplify the analysis. However, an
467 example model incorporating block effects is provided in the supplementary material and
468 accompanying analysis script. Season effects were also not modeled. Yield responses to foliar
469 fungicides in winter wheat have also been shown to depend on genotype and environment [46].
470 The model can be extended to incorporate blocks and/or season through the predictor or individual
471 curve parameters, with fixed or random effects specified according to how variation among groups
472 is modeled [47]. Prior distributions in this study were specified using information from published
473 studies and field trials. Expert knowledge provides an additional source of information for prior
474 specification and can be incorporated (obtained and used for prior specification) through structured
475 elicitation [48]. Similarly to what was done in this study, these priors informed by expert
476 knowledge can then be combined with observed data through Bayesian updating to obtain posterior
477 distributions [49, 50].

Conclusion

This study demonstrates a Bayesian LDPC framework for analyzing plant disease epidemics with a limited number of disease observations over time and limited information in the observations. By combining biological information (through prior distributions) with disease severity observations over time, the framework enables joint estimation of K, r, and y_0, and propagates parameter uncertainty (provided by our prior beliefs) into predictions and derived quantities. In CS1, posterior contrasts were utilized as a demonstration to compare posterior distributions for model parameters and derived quantities, and probabilistic inference was used to support conclusions of difference between treatments. In CS2, the posterior distribution of K closely resembled its prior, illustrating the dependence of inference on prior knowledge when observations provide little information about a parameter, and reinforcing the importance of a thoughtful prior specification. In this case, obtaining posterior estimates did not imply that all parameters had been identified from the observed data. Posterior predictions also incorporated observation-level variability, and therefore, prior uncertainty is just a partial contributor to predictive uncertainty. Limited predictive transfer across seasons highlighted the importance of model specification and predictive evaluation. This framework demonstrates how Bayesian inference can be harnessed as an additional tool for characterizing and comparing plant disease epidemics.

Abbreviations

AUDPC area under the disease progress curve

AUDPS area under the disease progress stairs

sAUDPC standardized area under the disease progress curve

sAUDPS standardized area under the disease progress stairs

LDPC logistic disease progress curve

FGS Feekes growth stage

PPI Posterior predictive interval

CrI Credible interval

CS case study

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Data availability The data and associated R code used to implement the models and perform the analyses are publicly available in the GitHub repository: https://github.com/ClaudioDSJr11/Bayesian-Logistic-Disease-Progress-Curve-Paper.

Competing interests We declare that the authors have no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Funding Not applicable.

Authors' contributions All authors (C.D., A.K., E.D., T.H. and K.A.O.) were involved in project conceptualization and manuscript review, writing, editing (all drafts). C.D., K.A.O, and E.D. were involved in data acquisition. C.D., T.H., and A.K. conducted formal analysis and validation. C.D., T.H. and K.A.O. were involved in the design of methodology. K.A.O. and T.H. provided supervision and project administration. K.A.O. provided resources and funding acquisition.

Acknowledgements The authors acknowledge the contributions of members of the Kansas State University wheat pathology lab for their role in data collection. The authors also acknowledge the Kansas Agricultural Experiment Station.

References

1. Miguez F, Archontoulis S, Dokoohaki H. Nonlinear regression models and applications. In: Glaz B, Yeater KM, editors. Applied statistics in agricultural, biological, and environmental sciences. 2018. doi:10.2134/appliedstatistics.2016.0003.c15.
2. Campbell CL. Disease progress in time: modelling and data analysis. In: Jones DG, editor. The epidemiology of plant diseases. Dordrecht: Springer; 1998. p. 181-206.
3. Jeger MJ. Analysis of disease progress as a basis for evaluating disease management practices. Annu Rev Phytopathol. 2004;42:61-82. doi: 10.1146/annurev.phyto.42.040803.140427.
4. Madden LV, Hughes G, van den Bosch F. Temporal analysis I: quantifying and comparing epidemics. In: Madden LV, Hughes G, van den Bosch F, editors. The study of plant disease epidemics. St. Paul: American Phytopathological Society; 2007. p. 63-116.
5. Segarra J, Jeger MJ, van den Bosch F. Epidemic dynamics and patterns of plant diseases. Phytopathology. 2001;91:1001-10. doi: 10.1094/PHYTO.2001.91.10.1001.
6. Van der Plank JE. Plant diseases: epidemics and control. New York: Academic Press; 1963.
7. Dougherty ER, Dalton LA, Alexander F. Small data is the problem. In: 2015 49th Asilomar Conference on Signals, Systems and Computers. 2015. p. 418-22. doi:10.1109/ACSSC.2015.7421161.
8. Dou B, Zhu Z, Merkurjev E, Ke L, Chen L, Jiang J, et al. Machine learning methods for small data challenges in molecular science. Chem Rev. 2023;123:8736-80. doi:10.1021/acs.chemrev.3c00189.
9. Ultsch A, Lötsch J. Augmenting small biomedical datasets using generative AI methods based on self-organizing neural networks. Brief Bioinform. 2025;26:bbae640. doi:10.1093/bib/bbae640.
10. Cobelli C, DiStefano JJ III. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. Am J Physiol. 1980;239:R7-24. doi:10.1152/ajpregu.1980.239.1.R7.
11. Preston SP, Wilkinson RD, Clayton RH, Chappell MJ, Mirams GR. Think before you fit: parameter identifiability, sensitivity and uncertainty in systems biology models. Curr Opin Syst Biol. 2025;100563. doi: 10.1016/j.coisb.2025.100563.
12. Cole D. Parameter redundancy and identifiability. Boca Raton: CRC Press; 2021. doi: 10.1007/s13253-021-00441-7.
13. Apgar JF, Witmer DK, White FM, Tidor B. Sloppy models, parameter uncertainty, and the role of experimental design. Mol Biosyst. 2010;6:1890-900. doi:10.1039/B918098B.
14. Campbell CL, Madden LV. Introduction to plant disease epidemiology. New York: John Wiley & Sons; 1990.
15. Neher DA, Campbell CL. Underestimation of disease progress rates with the logistic, monomolecular, and Gompertz models when maximum disease intensity is less than 100 percent. Phytopathology. 1992;82:811-4.
16. Kranz J. Comparative anatomy of epidemics. In: Horsfall JG, Cowling EB, editors. Plant disease: an advanced treatise. Vol. 2, How disease develops in populations. New York: Academic Press; 1978. p. 33-62.

17. Kranz J, editor. Epidemics of plant diseases: mathematical analysis and modeling. 2nd ed. Berlin: Springer-Verlag; 1990.
18. Hooten MB, Hefley TJ. Bringing Bayesian models to life. Boca Raton: CRC Press; 2019.
19. Semochkina D, Walsh C. Incorporating additional evidence as prior information to resolve non-identifiability in Bayesian disease model calibration: a tutorial. Stat Med. 2025;44. doi:10.1002/sim.70039.
20. Smithson M, Verkuilen J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychol Methods. 2006;11:54-71.
21. Ferrari S, Cribari-Neto F. Beta regression for modelling rates and proportions. J Appl Stat. 2004;31:799-815.
22. Kranz J. A study on maximum severity in plant diseases. In: Travaux dédiés à Georges Viennot-Bourgin. Paris: Société Française de Phytopathologie; 1977. p. 169-73.
23. Madden LV, Pirone TP, Raccah B. Temporal analysis of two viruses increasing in the same tobacco fields. Phytopathology. 1987;77:974-80.
24. Shearer BL, Crane CE, Barrett S, Cochrane A. Assessment of threatened flora susceptibility to Phytophthora cinnamomi by analysis of disease progress curves in shadehouse and natural environments. Australas Plant Pathol. 2007;36:609-20. doi:10.1071/AP07074.
25. Naseri B, Kazemi H. Structural characterization of stripe rust progress in wheat crops sown at different planting dates. Heliyon. 2020;6:e05328. doi:10.1016/j.heliyon.2020.e05328.
26. Omara R, Aly A, Abou-Zeid M. Characterization of partial resistance to stripe rust (Puccinia striiformis f. sp. tritici) in some Egyptian wheat cultivars. J Plant Prot Pathol. 2018;9:111-9. doi:10.21608/jppp.2018.41259.
27. Mabrouk O, Fahim M, Badeea OAE, Omara R. The impact of wheat yellow rust on quantitative and qualitative grain yield losses under Egyptian field conditions. Egypt J Phytopathol. 2022. doi:10.21608/ejp.2022.117996.1054.
28. De Rossi RL, Guerra FA, Plazas M, Vuletic E, Guerra G, Reis EM. Temporal analysis of northern corn leaf blight (Exserohilum turcicum Pass. Leonard & Suggs) epidemics. Rev Fac Cienc Agrar UNCuyo. 2025. doi:10.48162/rev.39.176.
29. Plummer M. rjags: Bayesian graphical models using MCMC. R package version 4-16. 2024. https://CRAN.R-project.org/package=rjags. Accessed 14 Oct 2026.
30. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2024. https://www.R-project.org/. Accessed 14 Oct 2026.
31. Conn PB, Johnson DS, Williams PJ, Melin SR, Hooten MB. A guide to Bayesian model checking for ecologists. Ecol Monogr. 2018;88:526-42. doi:10.1002/ecm.1314.
32. Pfender WF. Monocyclic and polycyclic root diseases: distinguishing between the nature of the disease cycle and the shape of the disease progress curve. Phytopathology. 1982;72:31. doi:10.1094/Phyto-72-31.
33. Carmona M, Sautua F, Pérez-Hernández O, Reis EM. Role of fungicide applications on the integrated management of wheat stripe rust. Front Plant Sci. 2020;11:733. doi:10.3389/fpls.2020.00733.

34. De Wolf E, Bockus W, Shoup D, Eddy R, Duncan S, Shroyer J. Evaluating the need for wheat foliar fungicides. Manhattan: Kansas State University Agricultural Experiment Station and Cooperative Extension Service; 2012. https://www.bookstore.ksre.ksu.edu/pubs/MF3057.pdf. Accessed 17 Dec 2024.
35. Andersen KF, Morris L, Derksen RC, Madden LV, Paul PA. Rainfastness of prothioconazole + tebuconazole for Fusarium head blight and deoxynivalenol management in soft red winter wheat. Plant Dis. 2014;98(10):1398-406. doi:10.1094/PDIS-01-14-0092-RE.
36. Gilligan CA. Comparison of disease progress curves.. The New phytologist. 1990;115 2: 223-242. doi:10.1111/j.1469-8137.1990.tb00448.x
37. Holb I, Heijne B, Gáll J, Jeger M. Analysis of summer epidemic progress of apple scab at different apple production systems in the Netherlands and hungary. Phytopathology. 2005;95 9: 1001-20 . doi:10.1094/phyto-95-1001
38. Jang G, Candan K, Chowell G. A comparative study of simulation-based inference methods for epidemic models with identifiability considerations. PLOS Computational Biology. 2026;22. doi:10.1371/journal.pcbi.1014364
39. Chowell G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A Primer for parameter uncertainty, identifiability, and forecasts. Infectious Disease Modelling. 2017;2:379 - 398. doi:10.1016/j.idm.2017.08.001
40. Gelman A, Shalizi CR. Philosophy and the practice of Bayesian statistics. Br J Math Stat Psychol. 2013;66:8-38. doi:10.1111/j.2044-8317.2011.02037.x.
41. Mikkola P, Martin OA, Chandramouli S, Hartmann M, Abril Pla O, Thomas O, et al. Prior knowledge elicitation: the past, present, and future. Bayesian Anal. 2024;19:1129-61. doi:10.1214/23-BA1381.
42. Tanner MA. Tools for statistical inference. 3rd ed. New York: Springer; 1991.
43. Cook JD, Williams DM, Walsh DP, Hefley TJ. Bayesian forecasting of disease spread with little or no local data. Sci Rep. 2023;13:8137. doi:10.1038/s41598-023-35177-6.
44. Chandran J, Gopinath PP, Pramod R, Vijayan A. Modeling plant disease epidemics: a comprehensive review of disease progress curves. J Sustain Technol Agric. 2025;1(2). doi:10.65287/josta.202512.72CB.
45. Alves KS, Del Ponte EM. Analysis and simulation of plant disease progress curves in R: introducing the epifitter package. Phytopathol Res. 2021;3:22. doi:10.1186/s42483-021-00098-7.
46. Cruppe G, DeWolf E, Jaenisch BR, Andersen Onofre K, Valent B, Fritz AK, et al. Experimental and producer-reported data quantify the value of foliar fungicide to winter wheat and its dependency on genotype and environment in the U.S. central Great Plains. Field Crops Res. 2021;273:108300. doi:10.1016/j.fcr.2021.108300.
47. Gelman A. Analysis of variance—why it is more important than ever. Ann Stat. 2005;33:1-53. doi:10.1214/009053604000001048.
48. Robledo J, Plex Sulá AI, Jaworski LG, Mouafo-Tchinda RA, Andersen Onofre KF, Thomas-Sharma S, et al. Expert knowledge elicitation: accessing the big data in experts' brains. Phytopathology. 2025;115:1245-59. doi:10.1094/PHYTO-06-25-0220-FI.

49. Garrett KA, Madden LV, Hughes G, Pfender WF. New applications of statistical tools in plant pathology. Phytopathology. 2004;94:999-1003. doi:10.1094/PHYTO.2004.94.9.999.
50. Yuen JE, Hughes G. Bayesian analysis of plant disease prediction. Plant Pathol. 2002;51:407-12. doi:10.1046/j.0032-0862.2002.00741.x.
51. Simko I, Piepho HP. The area under the disease progress stairs: calculation, advantage, and application. Phytopathology. 2012;102:381-9. doi:10.1094/PHYTO-07-11-0216.

686 **Table 1 –** Fungicide treatments used in the field experiments for both cultivars.

Treatment/Product	Application timing	Active Ingredient(s)	Application rate (fl. oz)	Mode of Action
Untreated	-	-	-	-
Healthy check (Nexicor, Prosaro, Headline)	FGS[b] 7 *fb*[a] FGS 9 *fb* FGS 10.5.1	Refer to individual product rows	13 *fb* 6.5 *fb* 6	Refer to individual product rows
Headline	FGS 7 FGS 9	Pyraclostrobin 23.6%	9	FRAC[c] 11 (QoI)
Prosaro	FGS 7 FGS 9	Prothioconazole 19%, Tebuconazole 19%	8.2	FRAC 3 (DMI)
Topguard	FGS7 FGS9	Flutriafol 11.8%	14	FRAC 3 (DMI)
Tebustar	FGS7 FGS9	Tebuconazole 38.7%	4	FRAC 3 (DMI)
Nexicor	FGS7 FGS9	Fluxapyroxad 2.81%, Pyraclostrobin 18.76%, Propiconazole 11.73%	13	FRAC 7 (SDHI), FRAC 11, FRAC 3

687 [a]*fb* = Followed by;
688 [b]FGS = Feekes Growth Stage;
689 [c]FRAC = Fungicide Resistance Action Committee.
690
691
692
693
694
695
696
697
698
699
700
701

Table 2 – Derived quantities for summarizing and comparing logistic disease progress curves.

Derived quantity[a]	Equation	Units	Description
AUDPC	$\int_{t_1}^{t_n} \frac{K}{1+\left(\frac{K-y_0}{y_0}\right)e^{-rt}}dt =$ $\frac{K}{r}[\log(e^{rt_n}+A)-\log(e^{rt_1}+A)]$ $A=\left(\frac{K-y_0}{y_0}\right)$ t_n – Last point in time t_1 – First point in time Or $\sum_{i=1}^{n-1}\frac{y_i+y_{i+1}}{2}*(t_{i+1}-t_i)$	Proportion*time unit	Absolute (raw) area under the disease progress curve **[4, 6]**
sAUDPC	$\frac{AUDPC}{t_n-t_1}$	Proportion	Standardized AUDPC by the length of the epidemic **[13]**
AUDPS[b]	$AUDPC+\left(\frac{y_1+y_n}{2}\right)*\left(\frac{t_n-t_1}{n-1}\right)$	Proportion*time	Area under the disease progress stairs Attribute proper weight to end points of the curve **[51]**
sAUDPS	$\frac{AUDPS*(n-1)}{t_n-t_1}$	Proportion	Standardized AUDPS by the length of the epidemic **[51]**

$\boldsymbol{r_*}$	$\frac{Kr}{4}$	Proportion/time unit	Maximum instantaneous rate of growth (First derivative of the LDPC, at the inflection point ($\mu = K_*$))
$\boldsymbol{K_*}$	$\frac{K}{2}$	Proportion	Disease severity at the inflection point, where the instantaneous rate of growth is maximal (Second derivative of the LDPC = 0)
$\boldsymbol{t_y}$	$\frac{1}{r}\log\left(\frac{y(K-y_0)}{y_0(K-y)}\right)$	Time unit (e.g., Day)	Time at which epidemic reaches y (Rearrangement of the LDPC)
Δ	$\theta_1 - \theta_2$	Same unit as θ (e.g., Proportion for K)	Difference between posterior distribution draws of a parameter (θ)

[a]Derived quantities should be calculated for each posterior draw of the quantities involved on the equation;
[b]Where n represents the number of assessment times used to calculate the AUDPC.

Table 3 – Posterior contrasts between applications made at FGS 7 and FGS 9 in CS1.

Contrast[a]	Posterior mean	95% Credible Interval		$P(\Delta > 0)$
		Lower bound	Upper bound	
$\Delta = K_7 - K_9$	0.13	-0.12	0.47	0.81
$\Delta = r_7 - r_9$	0.12	0.06	0.19	1
$\Delta = AUDPC_7 - AUDPC_9$	16.51	11.50	20.64	1
$\Delta = r_{*7} - r_{*9}$	0.03	0.01	0.04	0.99

[a]All contrasts were made in the format FGS 7 – FGS 9;
[b]$P(\Delta > 0) = 1$ indicates all posterior draws from the contrast were positive; $P(\Delta > 0) = 0.5$ indicates posterior support of both groups is the same.

739 **Table 4** – Prior and posterior summaries of model parameters in CS2.

Parameter	Prior distribution			Posterior distribution		
	Prior mean	95% Prior Interval		Posterior mean	95% Credible Interval	
		Lower bound	Upper bound		Lower bound	Upper bound
$\boldsymbol{K}$	0.75	0.42	0.96	0.75	0.42	0.96
$\boldsymbol{r}$	0.25	0.01	0.49	0.016	$5x10^{-4}$	0.05
$\boldsymbol{y_0}$	0.05	0.003	0.097	0.013	0.004	0.026

740

741

742

743

744

745

746

747

748

749

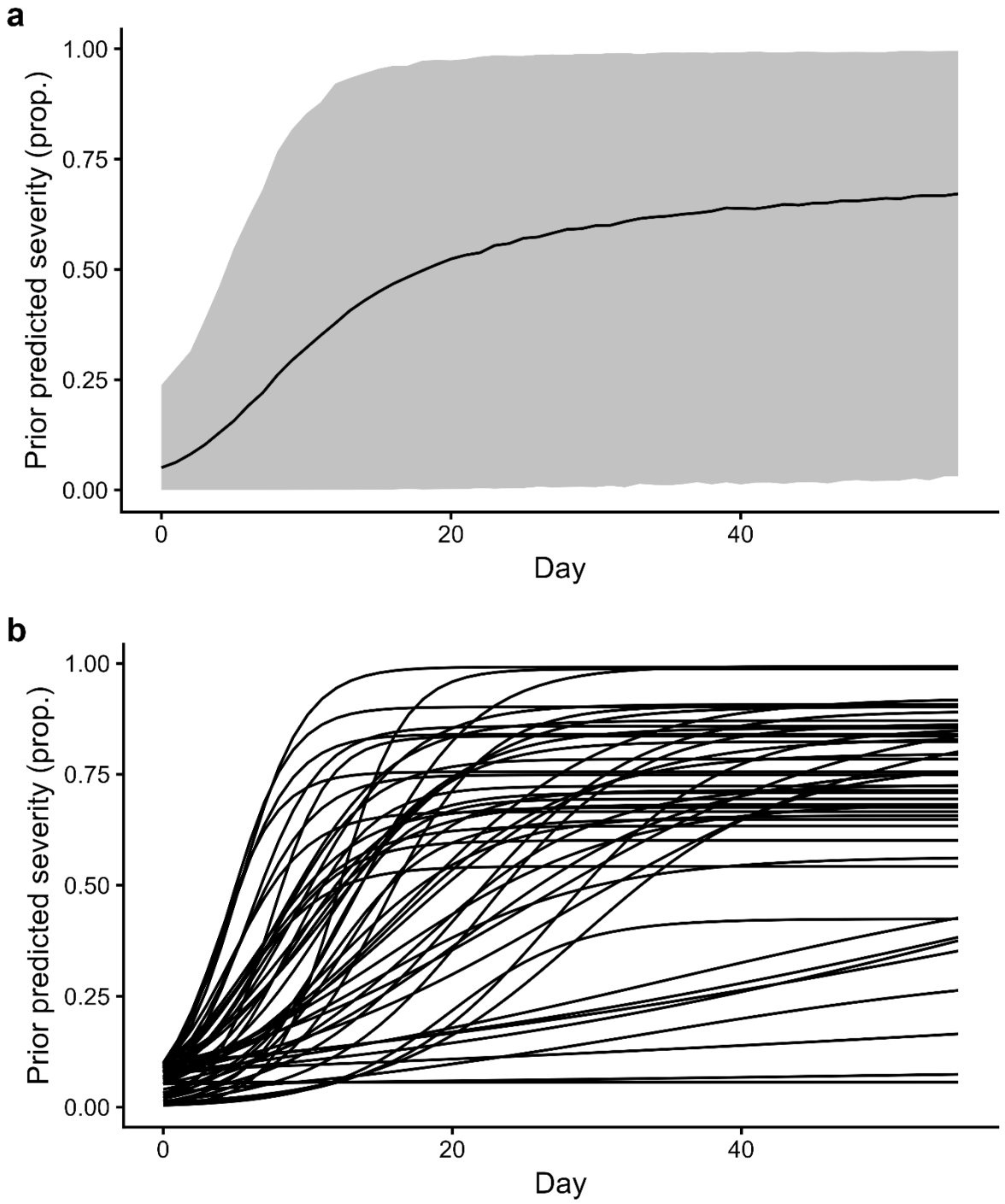


Figure 1 – a. Prior predictive mean disease severity in proportion (black line) and 95% prior predictive interval (gray ribbon) from simulated disease severity (y^*_{prior}). **b.** Fifty expected disease development trajectories calculated from randomly selected prior draws of μ_{prior} illustrating variation in these trajectories implied by the priors.

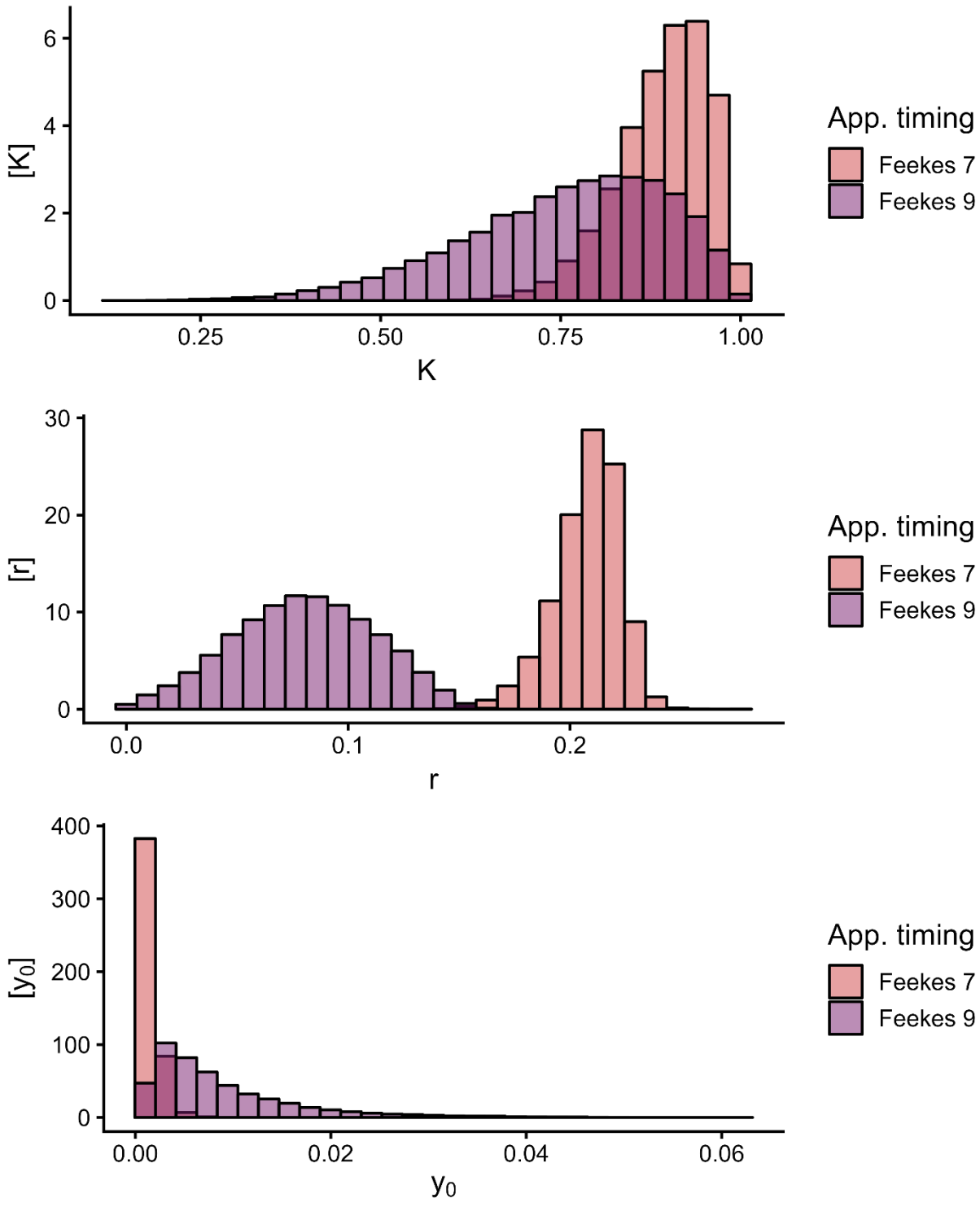
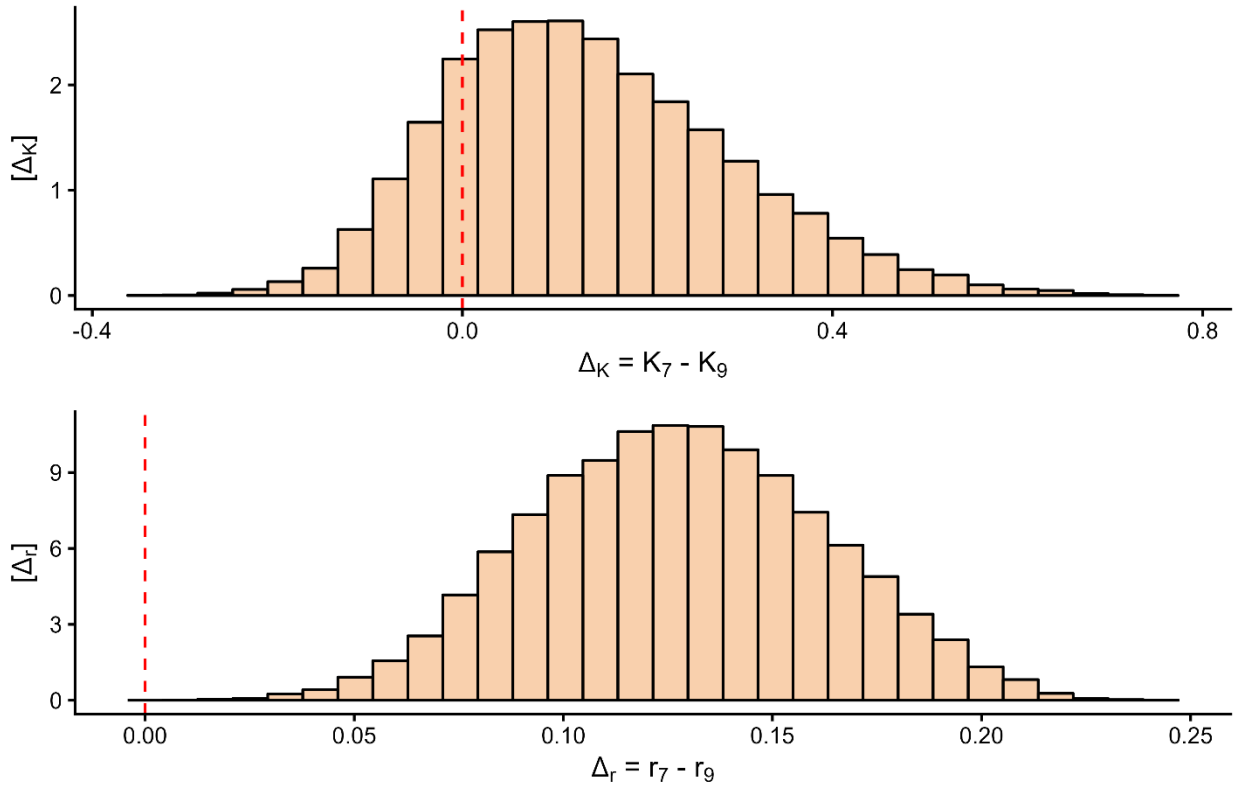
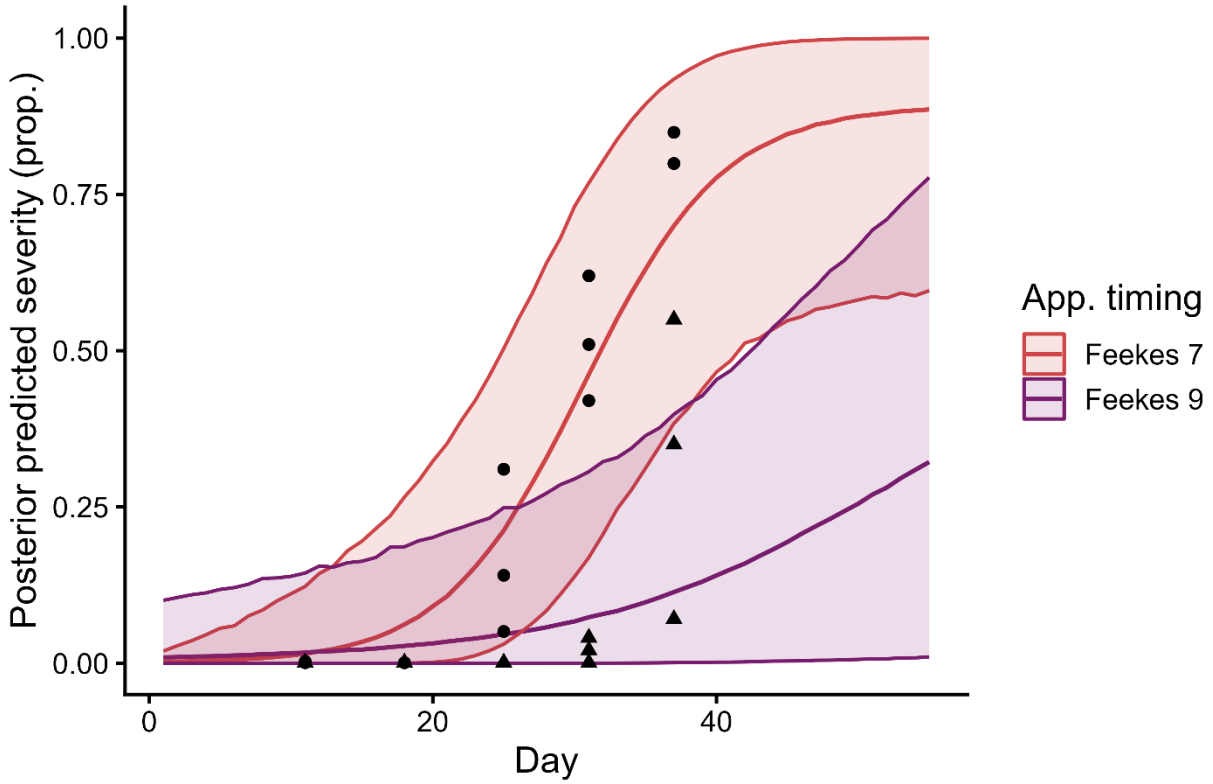


Figure 2 – Posterior distributions for the three logistic disease progress curve parameters K (upper), r (mid), and y_0 (lower) for the model fit to CS1 data. In the y-axis, density is denoted by "[]".



760

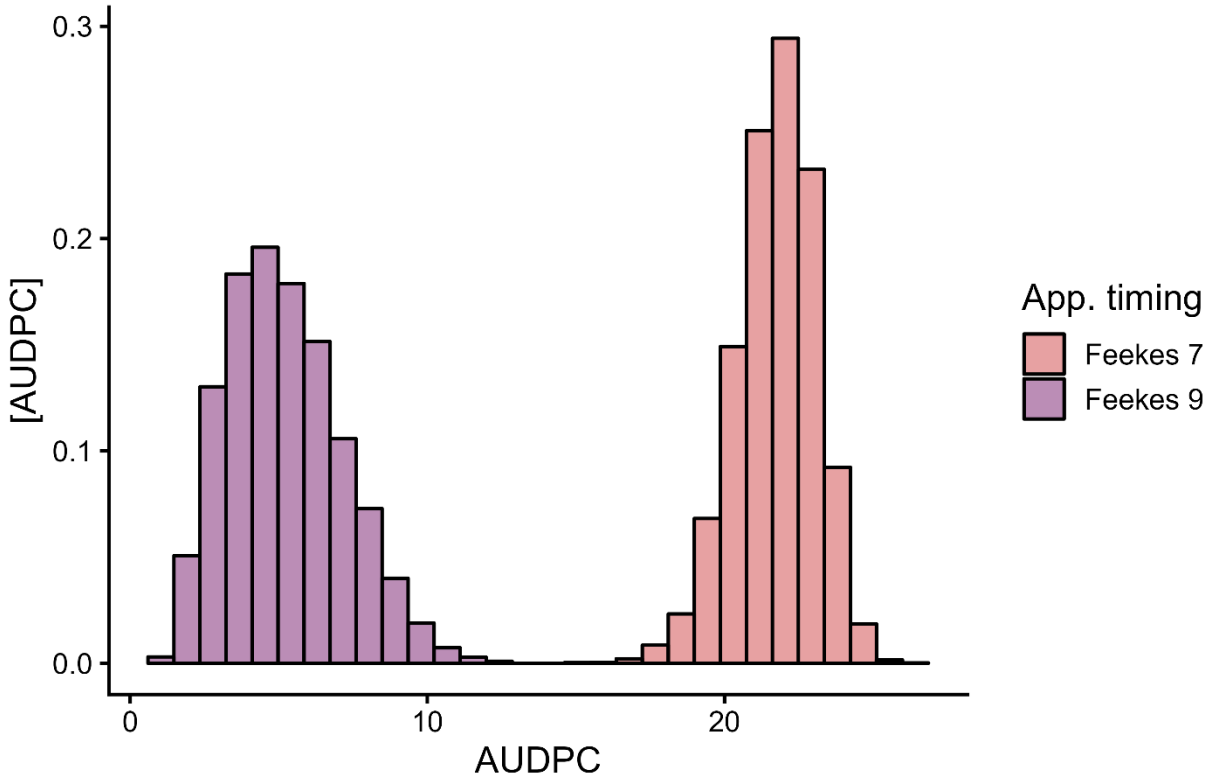
761 **Figure 3** – Posterior contrasts for K (upper) and r (lower) for the applications made at FGS 7
762 and at FGS 9. Dashed red line indicates zero to guide visualization of positive posterior draws. In
763 the y-axis, density is denoted by "[]".



764

765 **Figure 4** – Posterior predicted curves over a 55 days period with 95% PPI for both timings of
766 application (FGS 7 – orange, and FGS 9 – purple). Central line represents mean posterior curve,
767 and ribbon represents 95% PPI with external lines demarking its edges (upper and lower). Points
768 represent disease severity observations (y^*), with round shape representing FGS 7 and triangular
769 shape representing FGS 9.

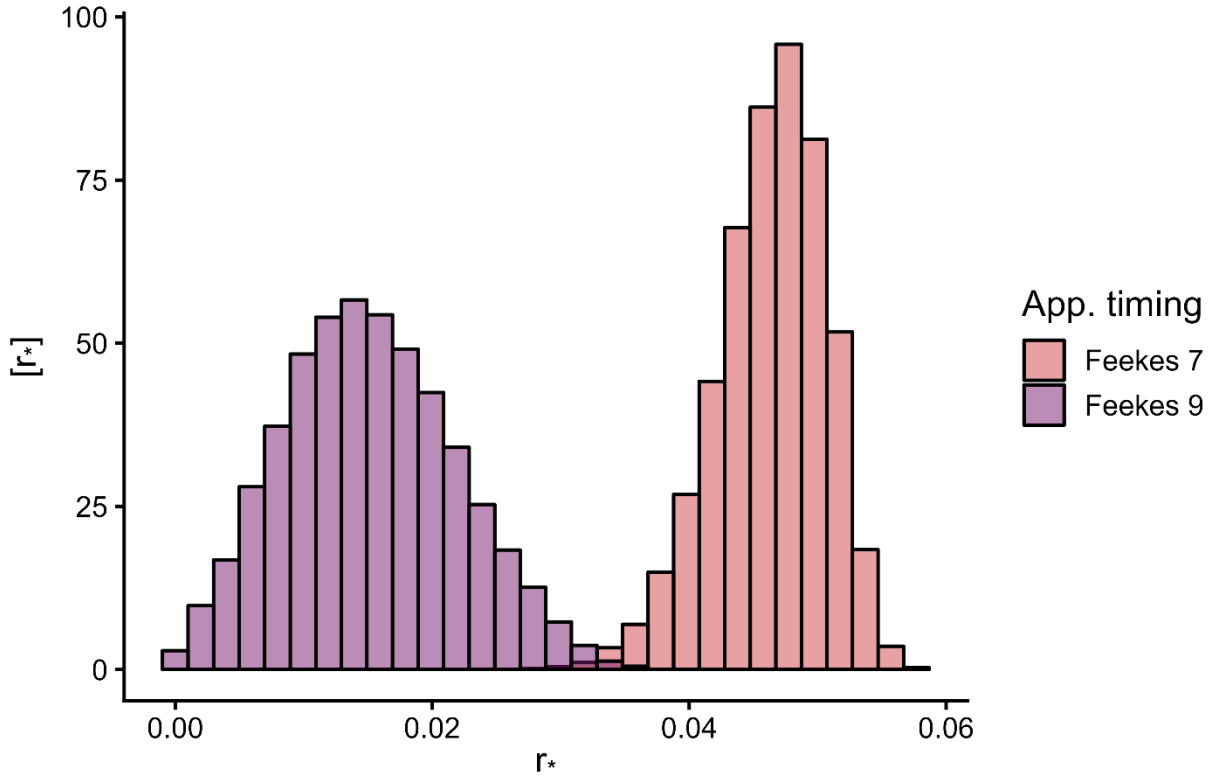
770



771

772 **Figure 5** – Posterior distributions for the derived quantity AUDPC for both timings of
773 application (FGS 7 – orange, and FGS 9 – purple). In the y-axis, density is denoted by "[]".

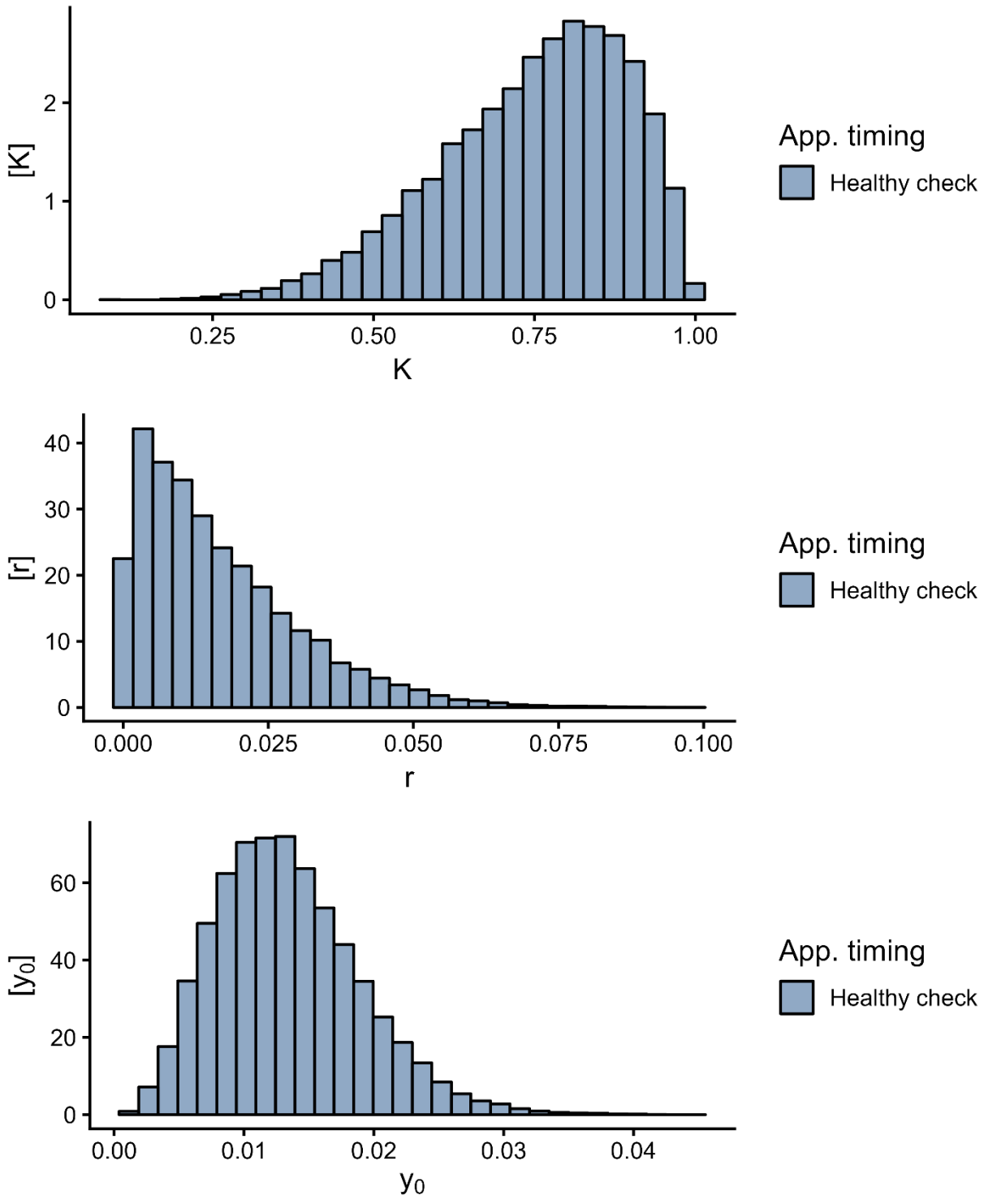
774



775

776 **Figure 6** – Posterior distributions for the derived quantity r_* for both timings of application (FGS
777 7 – orange, and FGS 9 – purple). In the y-axis, density is denoted by "[]".

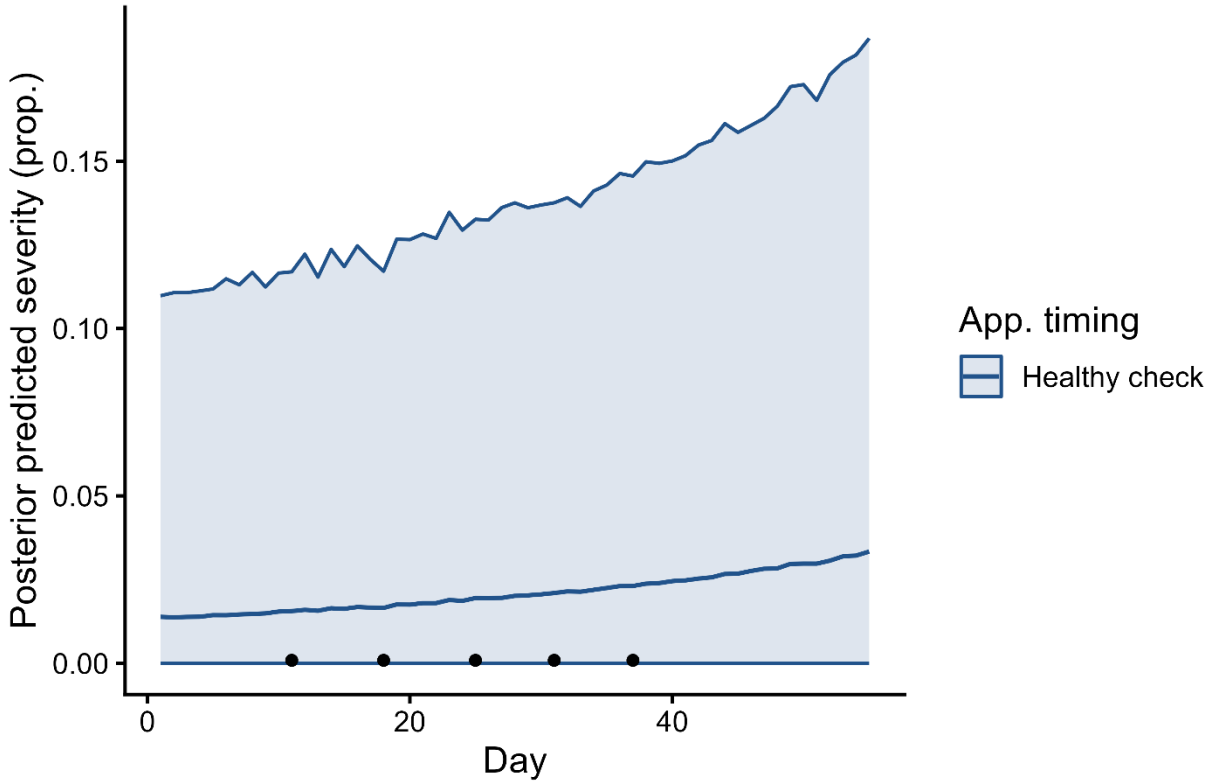
778



779

780 **Figure 7** – Posterior distributions for the three logistic disease progress curve parameters K
781 (upper), r (mid), and y_0 (lower) for the model fit to CS2 data. In the y-axis, density is denoted by
782 "[]".

783



784

785 **Figure 8** – Posterior predicted curve over a 55 days period with 95% PPI for the healthy check in
786 CS2. Central blue line represents mean posterior curve, and ribbon represents 95% PPI with
787 external lines demarking its edges (upper and lower). Points represent disease severity
788 observations (y^*).

789

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Sup.Material.docx