# A Novel Approximate Bayesian Inference Method for Compartmental Models in Epidemiology using Stan

University of St Andrews

***Joint Universities Pandemic and Epidemiological Research***
***JUNIPER Seminar Series***
***January 29th, 2025***


Xiahui Li
xl94@st-andrews.ac.uk
University of St Andrews

Dr Fergus Chadwick
fergusjchadwick@st-andrews.ac.uk
University of St Andrews

Dr Ben Swallow
bts3@st-andrews.ac.uk
University of St Andrews

# CONTENTS


1 Motivation and Background

2 Review Existing Methods

3 Introduce Proposed Method

4 Numerical Studies

5 Limitations and Improvements

# 1.1 Mechanistic Infectious Disease Models

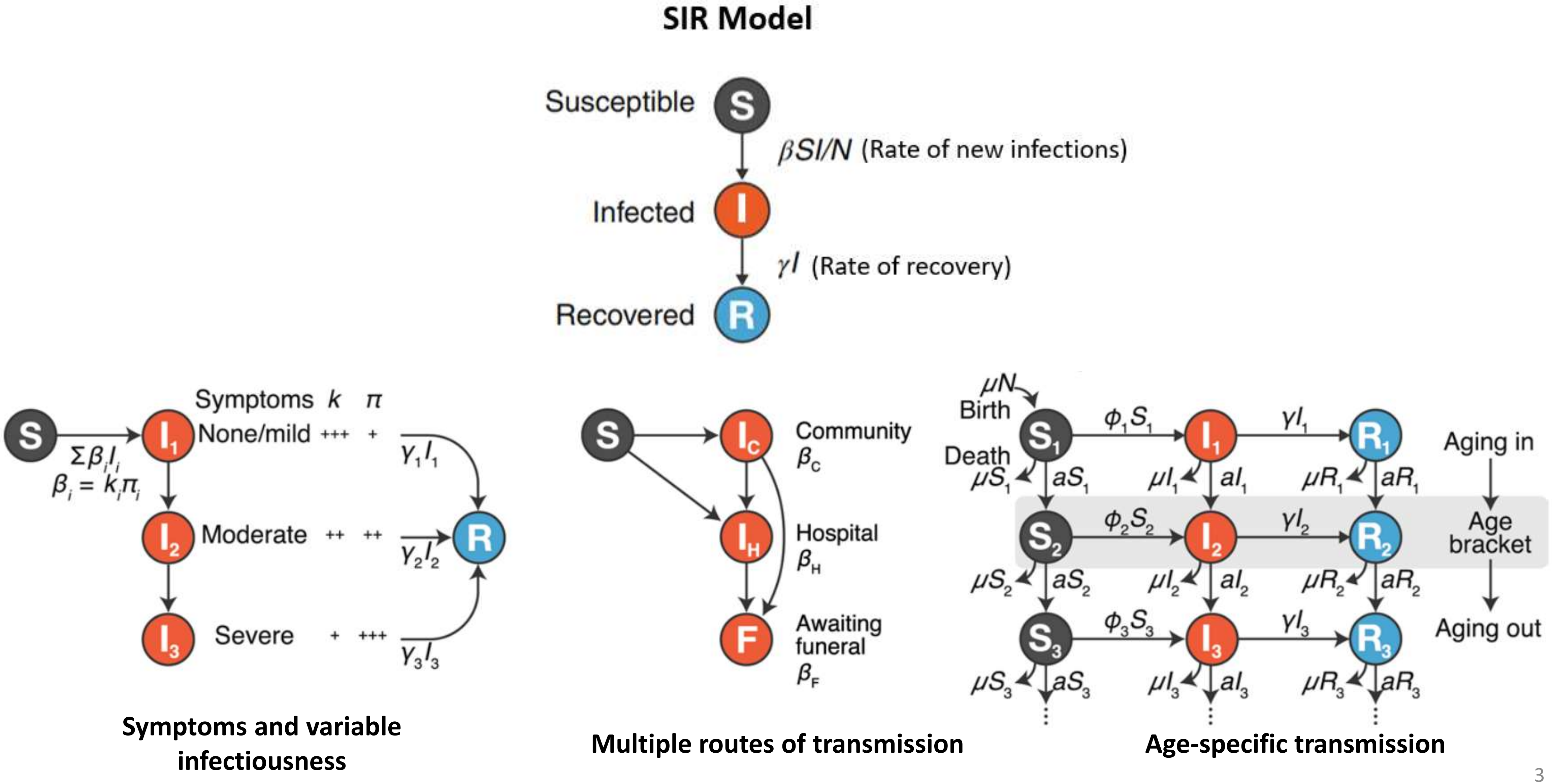
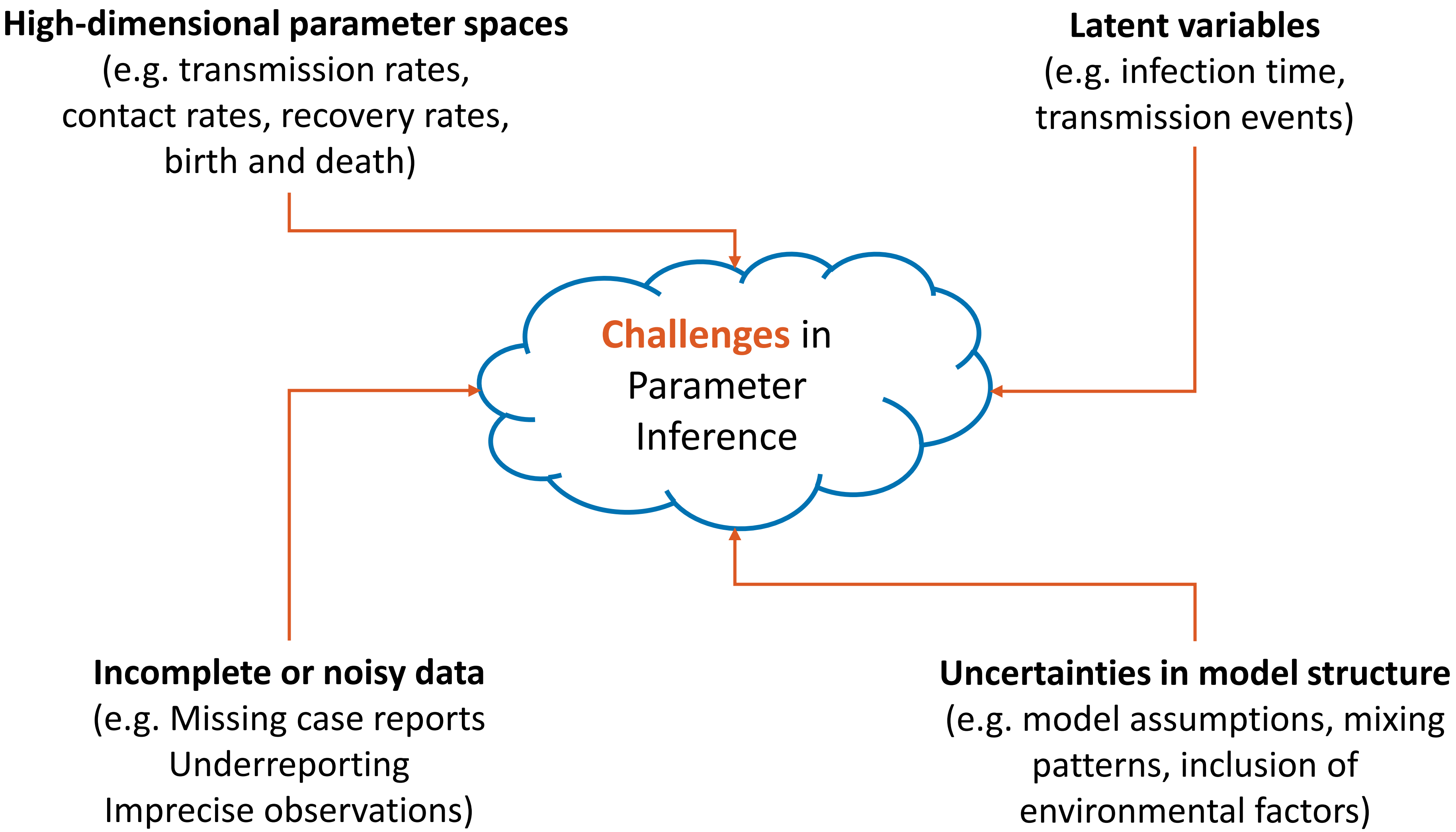





**Figure 1: SIR model extensions.** Adapted from “The SEIRS model for infectious disease dynamics,” by Bjørnstad, O. N., Shea, K., Krzywinski, M., & Altman, N. , 2020.

# 1.2 Challenges in Parameter Inference




Reference: Swallow, B., Birrell, P., Blake, J., Burgman, M., Challenor, P., Coffeng, L. E., ... & Vernon, I. (2022). Challenges in estimation, uncertainty quantification and elicitation for pandemic modelling. *Epidemics*, *38*, 100547.

# 1.2 Bayesian Statistical Inference

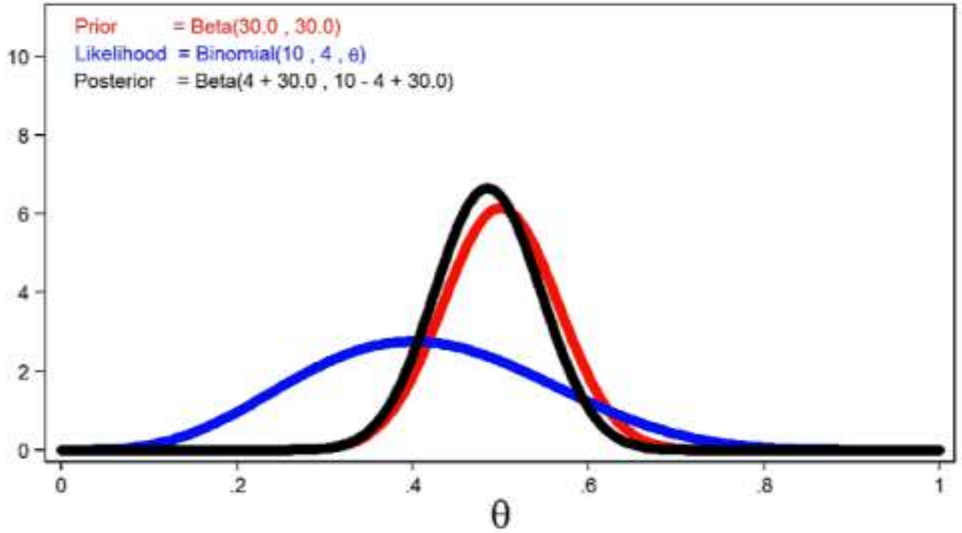
**Bayes' rule**

Likelihood

$$\pi(\theta|y_{obs}) \propto \pi(y_{obs}|\theta)\pi(\theta)$$

Posterior

Prior




GIF Animation 1: The effect of larger sample sizes on the posterior distribution.
Adapted from "Introduction to Bayesian statistics," by Chuck Huber, 2016.
https://blog.stata.com/2016/11/01/introduction-to-bayesian-statistics-part-1-the-basic-concepts/

# 1.3 Exact and Approximate Bayesian Method

Computationally Prohibitive

$$\boldsymbol{\pi(\theta|y_{obs}) \propto \pi(y_{obs}|\theta)P(\theta)}\ \textbf{in Complex Model}$$

Analytically Impossible

TRACTABLE Likelihood

INTRACTABLE Likelihood

## Exact Bayesian Method

Directly sample from the true posterior distribution

**Popular Examples**

Metropolis-Hastings Algorithm
Hamiltonian Monte Carlo (HMC)

## Approximate Bayesian Method

Bypass direct likelihood computation by matching simulated data to observations.

**Popular Examples**

Approximate Bayesian Computation

Synthetic Likelihood Method

# 2.1 Approximate Bayesian Computation (ABC) - Workflow

**Input**

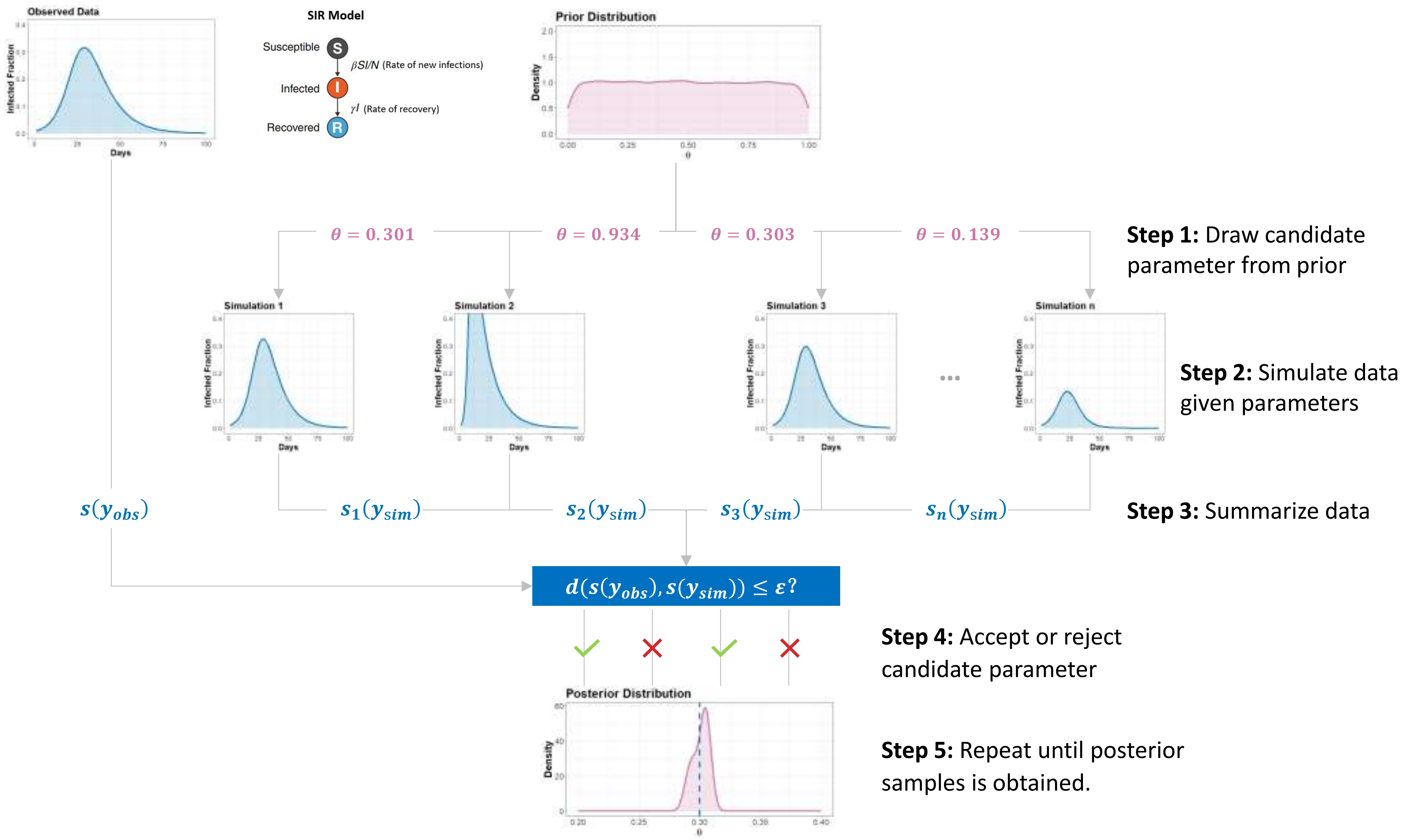
Observed data $y_{obs}$

Prior $\pi(\theta)$

Summary Statistics $s(\cdot)$

Distance $d(\cdot,\cdot)$

Tolerance $\varepsilon$



# 2.2 ABC Posterior Distribution

**Bayes' rule**

$$\boldsymbol{\pi(\theta|y_{obs}) \propto \pi(y_{obs}|\theta)\pi(\theta)}$$

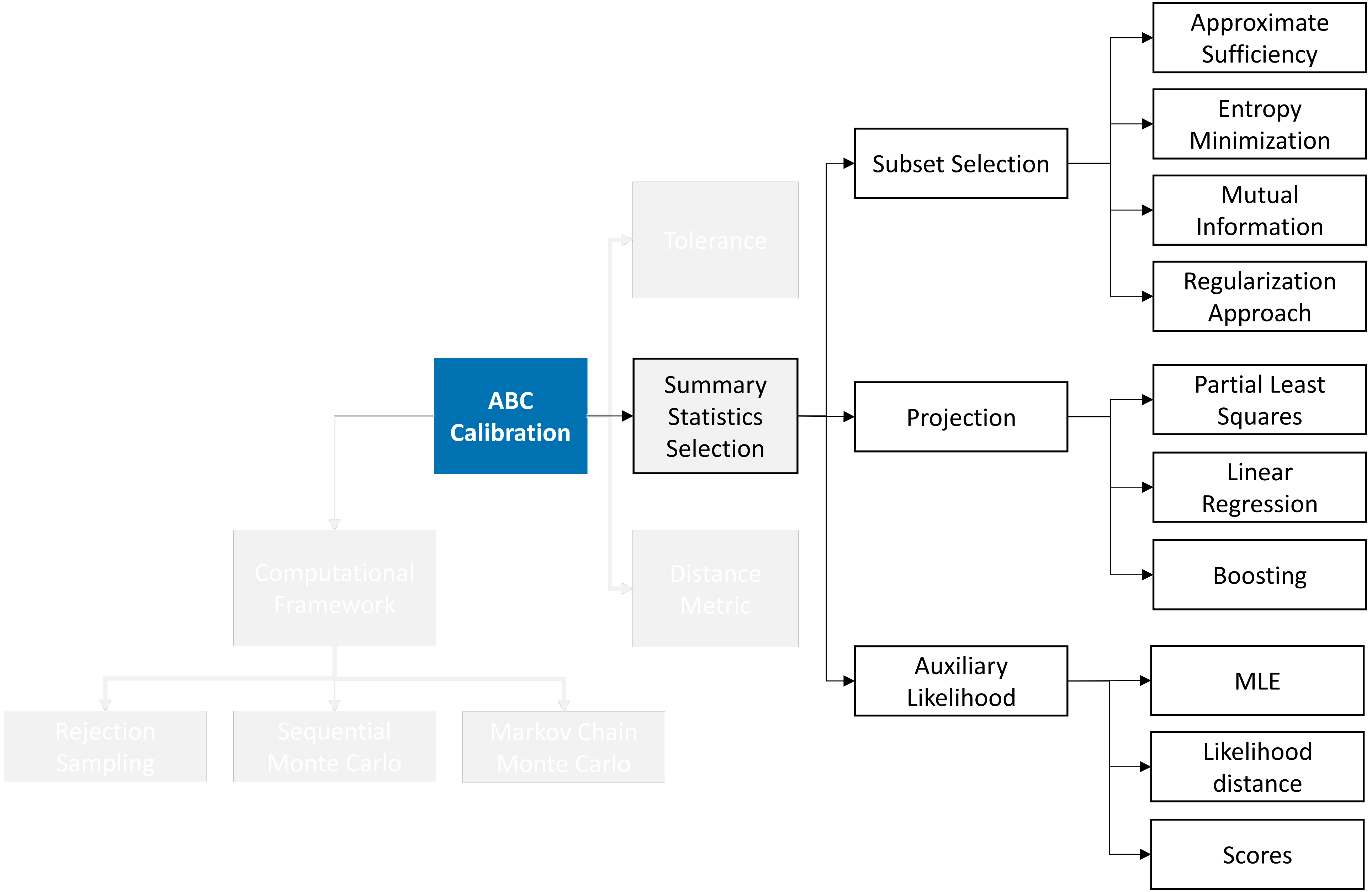
$$\boldsymbol{\pi_{ABC}(\theta, s(y_{sim})|s(y_{obs})) \propto K_h(\underbrace{d}_{\text{Distance}}(\underbrace{\boxed{s(y_{obs}), s(y_{sim})}}_{\text{Summary Statistics}}) \leq \underbrace{\varepsilon}_{\text{Tolerance}})\pi(s(y_{sim})|\theta)\pi(\theta)}$$

**Marginal distribution: integrate the joint distribution over synthetic data**

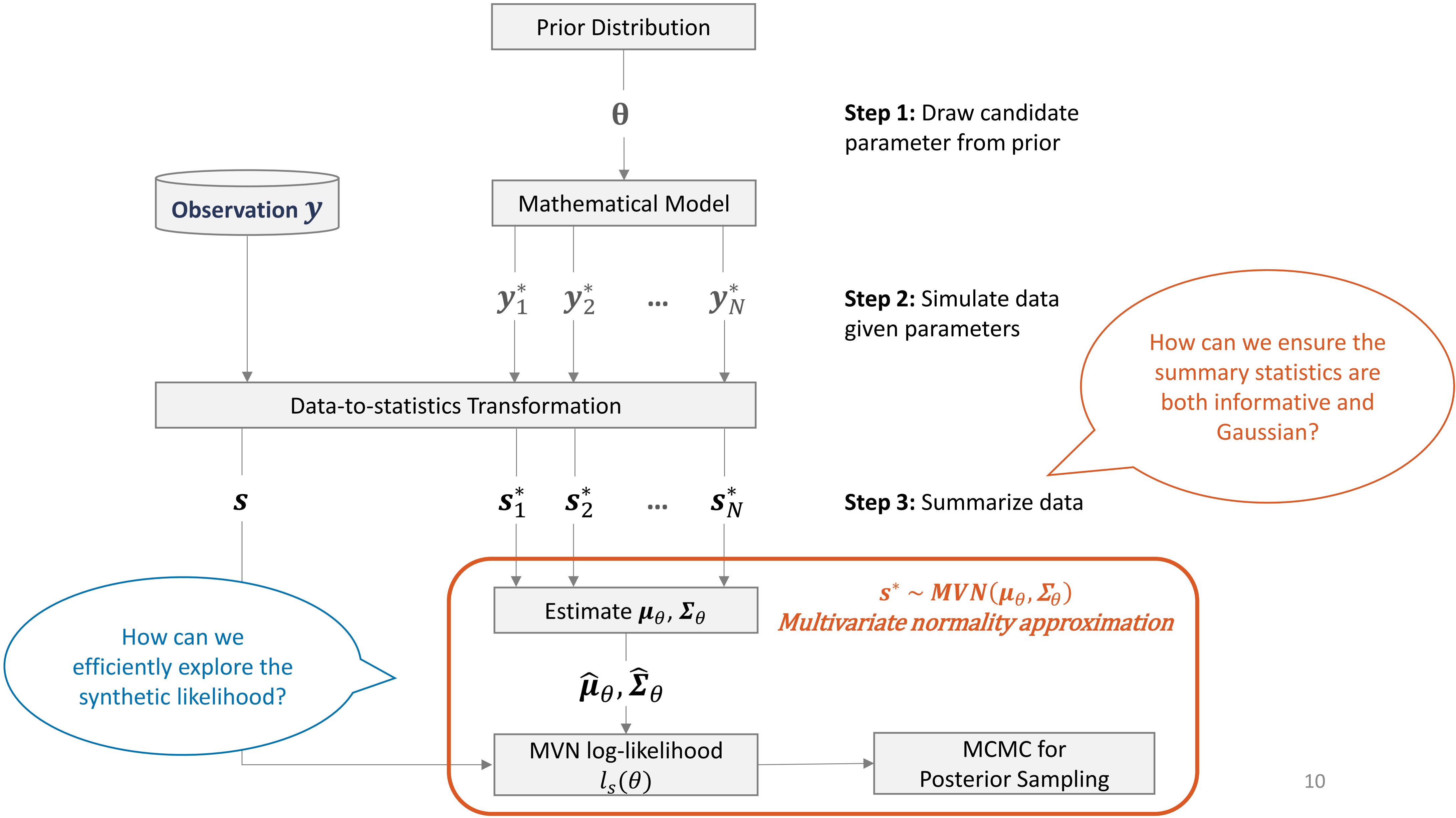
$$\pi_{ABC}(\theta|s(y_{obs})) \propto \int K_h\left(d(s(y_{obs}), s(y_{sim}))\right)\pi(s(y_{sim})|\theta)\pi(\theta)d(s(y_{sim}))$$

**ABC posterior distribution**

# 2.3 ABC Recent Advancement



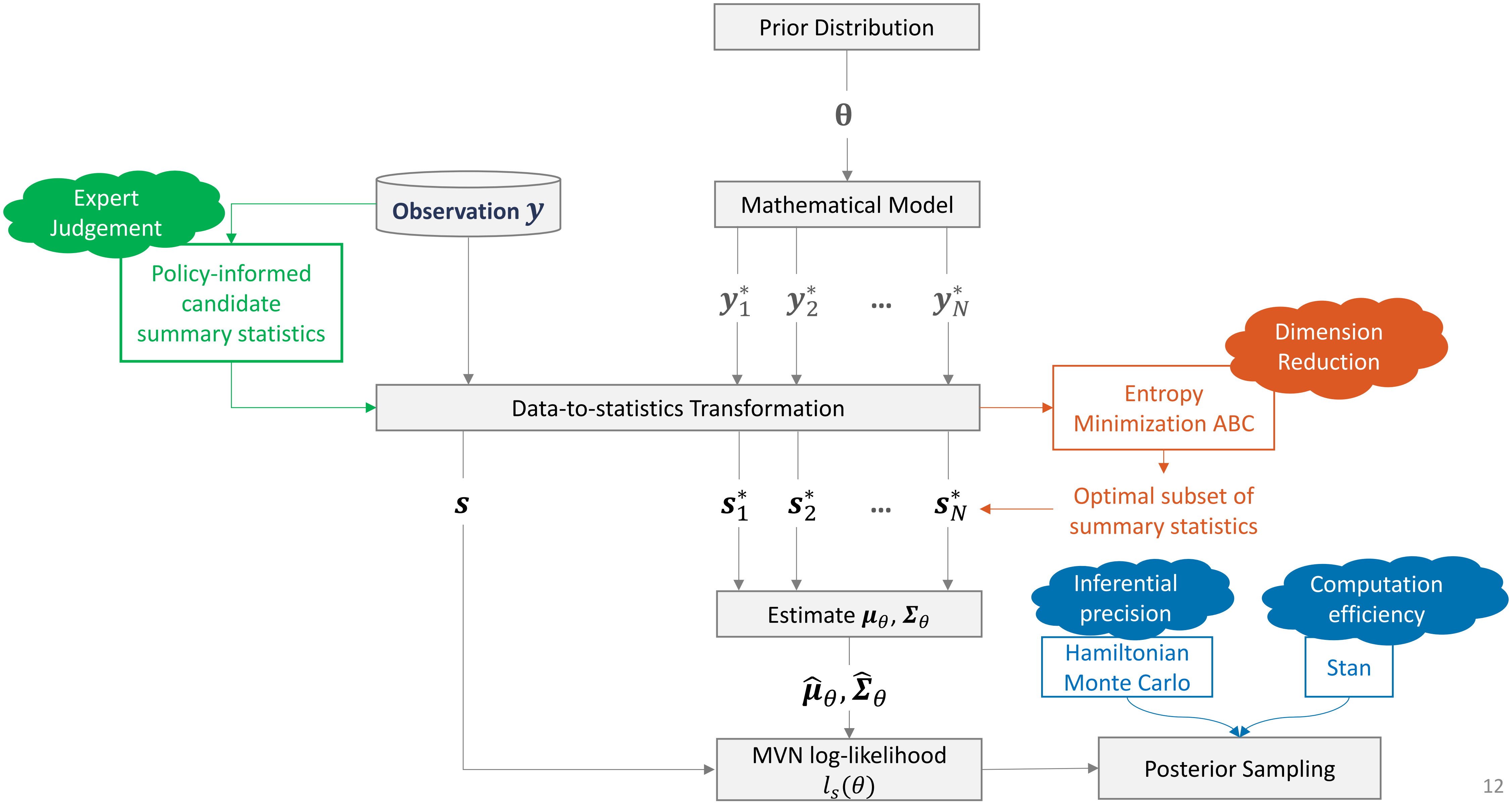
# 2.4 Synthetic Likelihood Method (Wood, 2010)



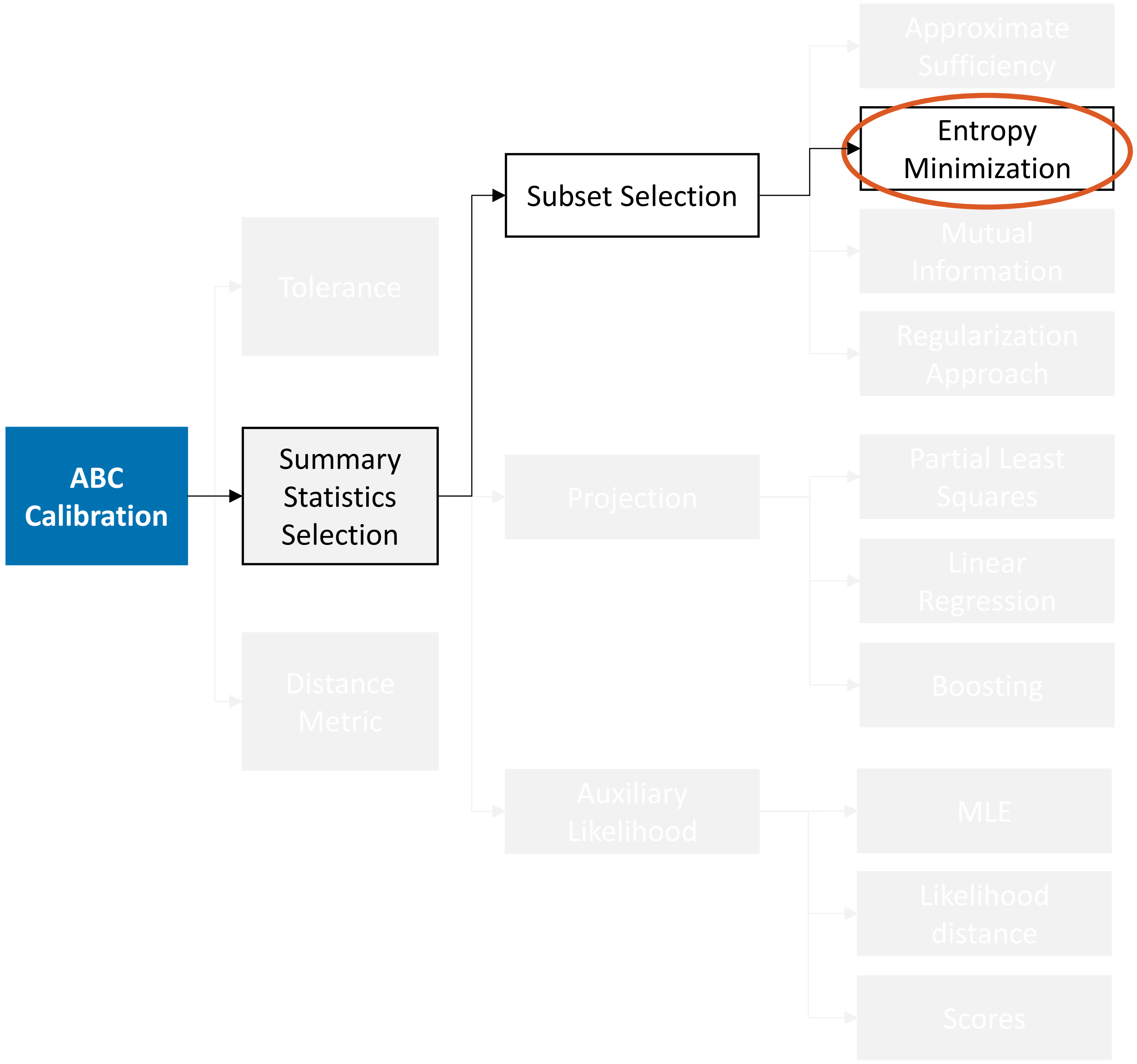
# 2.5 Method Comparison

| FEATURES | MCMC | ABC | BSL |
| --- | --- | --- | --- |
| Handle **intractable likelihoods**? | | ✓ | ✓ |
| **Likelihood-based** approximation? | ✓ | | ✓ |
| **Summary statistics** not required? | ✓ | | |
| **Computationally efficient** for high-dim or complex models? | | | ✓ |
| **Easy to interpret**? | ✓ | | ✓ |
| Avoid assuming **normality**? | ✓ | ✓ | |
| **Flexible** across different model types? | | ✓ | |

# 3.0 Proposed Method - Framework



# 3.1 Key Innovation 1: ABC for Summary Statistics Selection




## Entropy Minimisation

- Perform rejection-ABC and compute the $kth$ nearest neighbour of entropy on the ABC posterior sample.

$$\hat{E} = log\left[\frac{\pi^{p/2}}{\Gamma(\frac{p}{2}+1)}\right] - \Psi(\kappa) + \log(n) + \frac{p}{n}\sum_{i=1}^{n}\log(R_{i,\kappa})$$

- Removes redundant summaries with nonparametric entropy estimators, keeping only the most informative ones.

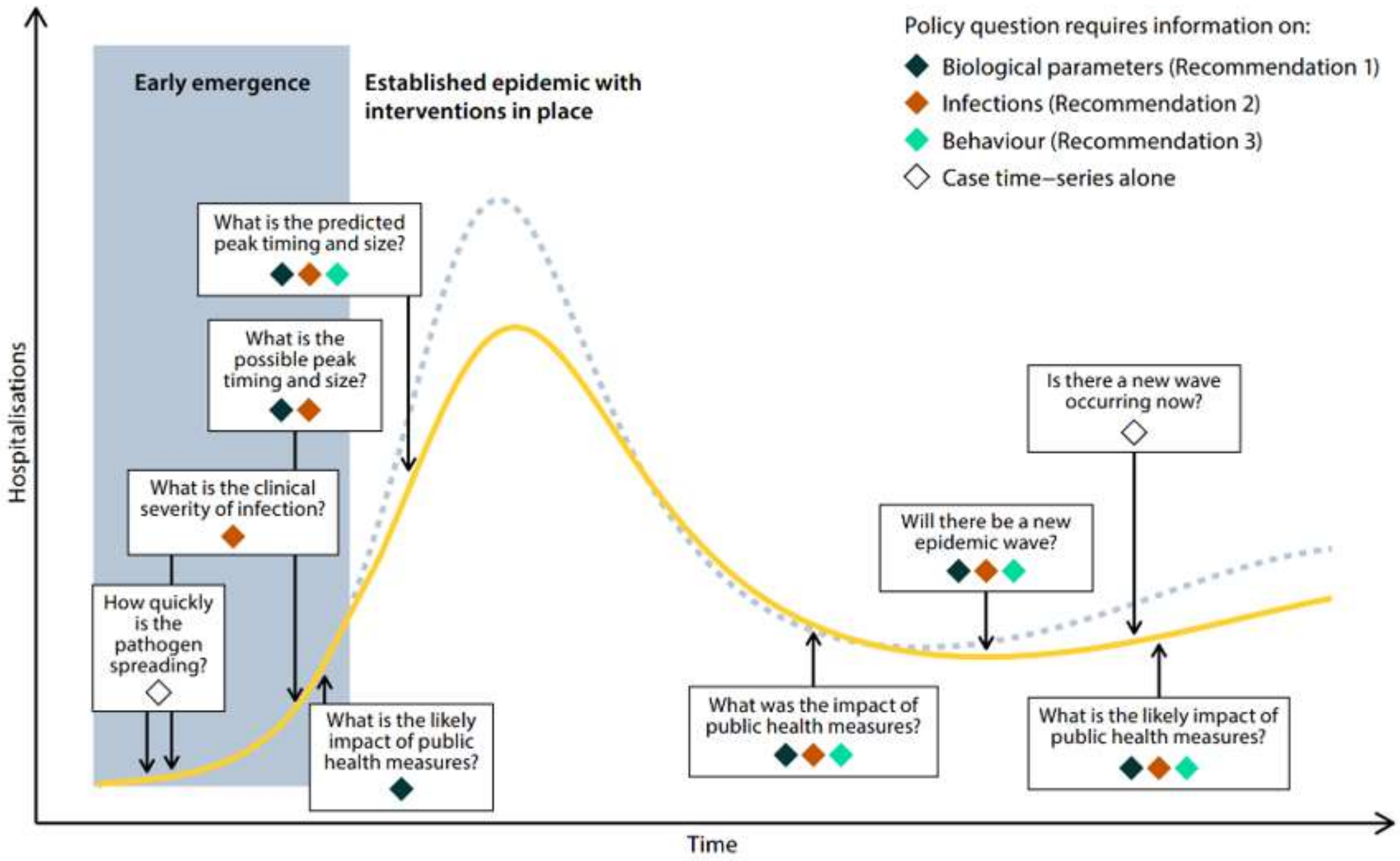
## Why this method?

- compared to more complex auto-tuning methods, entropy minimization provides a balance of simplicity, flexibility, and interpretability—especially when data are limited.

# 3.2 Key Innovation 2:  Policy-informed Candidate Statistics

- Real world epidemics are strongly influenced by external factors such as public health guidelines and intervention strategies.
- Including policy-driven metrics in summary statistics ensures both statistical value and practical relevance for decision-making.

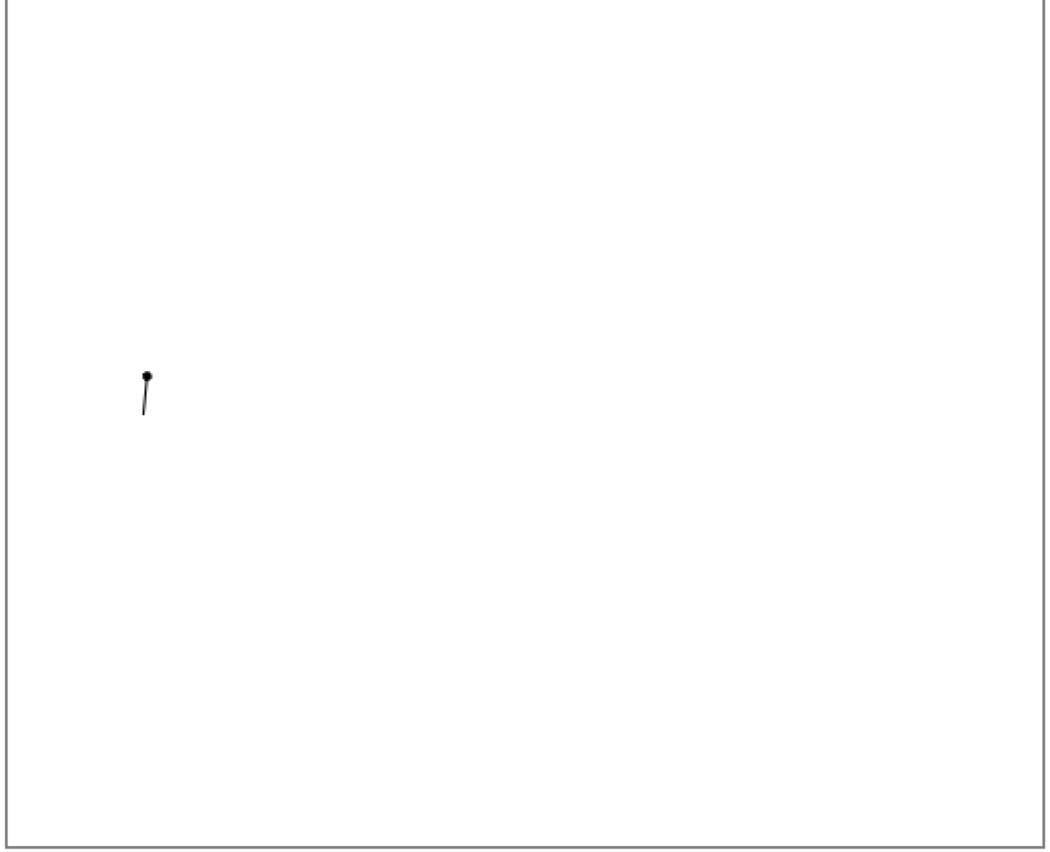
**Figure 1: Exemplar policy questions and transmission-related surveillance needs**




**Data Sources for Strategic Decision-making**. Adapted from “Opportunities to strengthen respiratory virus surveillance systems in Australia: Lessons learned from the COVID-19 response”, by Shearer, F. M., Edwards, L., Kirk, M., Eales, O., Golding, N., Hassall, J., ... & McCaw, J. M. ,2024.

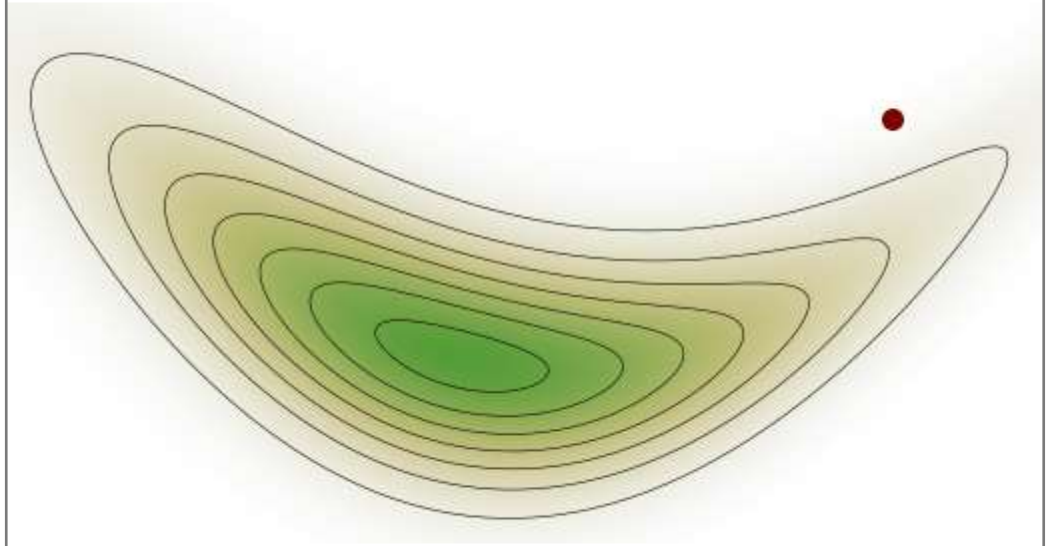
# 3.3 Key Innovation 3:  Compare MCMC Methods

**Random Walk Metropolis Hastings**



GIF Animation 1: Random walk Metropolis Hastings
Adapted from
https://bookdown.org/danbarch/psy_207_advanced_stats_I/MCMC-methods.html

**Hamiltonian Monte Carlo**



GIF Animation 2: Hamiltonian Monte Carlo sampling
Adapted from Justinkunimune - Own work
using: github.com/jkunimune/hamiltonian-mc, CC0

# 3.3 Key Innovation 3: Effective Likelihood Exploration via HMC

**Step 1: Initialization**

Specify likelihood $\pi(y|\theta)$ and prior $\pi(\theta)$

Derive posterior distribution $\pi(\theta|y)$

Derive gradient of log posterior $\nabla_\theta \log \pi(\theta|y)$

Simulate momentum $r \leftarrow N(0, M)$

Set starting position for leapfrog $\theta' \leftarrow \theta^{(t-1)}$ $r' \leftarrow r$

**Step 2: Leapfrog Integration**

Define joint density $\pi(\theta, r) = \pi(\theta|y) \cdot N(r|0, M)$

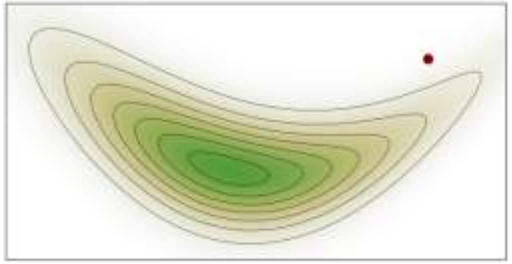
**Run leapfrog algorithm L times**

Solve Hamiltonian equation

$H(\theta, r) = -log\pi(\theta, r)$

Change of location: $\frac{d\theta}{dt} = \frac{\partial H}{\partial r}$

Change of momentum: $\frac{dr}{dt} = \frac{\partial H}{\partial \theta}$



Produce HMC proposal $(\theta', r')$

**Step 3: Metropolis Acceptance**

Accept $(\theta', r')$ with probability $\alpha = \min(1, \frac{\pi(\theta', r')}{\pi(\theta, r)})$

✓ $\theta^t \leftarrow \theta'$ $r^t \leftarrow -r'$

✗ $\theta^t \leftarrow \theta^{(t-1)}$ $r^t \leftarrow r^{(t-1)}$

$N$ simulations?

✓ Simulation Complete $\theta = (\theta^1, \dots, \theta^N)$

✗ (return to Simulate momentum)

GIF Animation 2: Hamiltonian Monte Carlo sampling a two-dimensional probability distribution.
Adapted from Justinkunimune - Own work using: github.com/jkunimune/hamiltonian-mc, CC0

# 3.3 Key Innovation 3: HMC using Stan

## Stan

- An **open-source platform** for high-performance **statistical modelling and computation**.
- Stan is a **probabilistic programming language** enabling
  - Compilation of models into efficient C++ for faster sampling
  - Parameter estimation via HMC, solving DE, and convergence checks.
  - Output of posterior samples and key quantities for analysis.

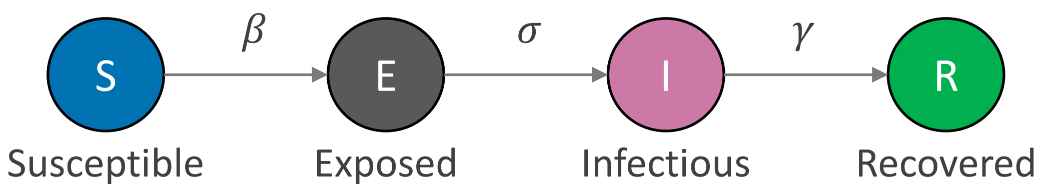
**Stan Model**

- functions
- data
- transformed data
- parameters
- transformed parameters
- model
- generated quantities

# 4.1 A Simulation Study

**Objective:** Estimate the SEIR parameters using proposed framework.

**Deterministic SEIR Model**
**(A simulated epidemic)**




**Ordinary Differential Equations**

$$\frac{dS}{dt} = -\frac{\beta SI}{N}$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E$$

$$\frac{dI}{dt} = \sigma E - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

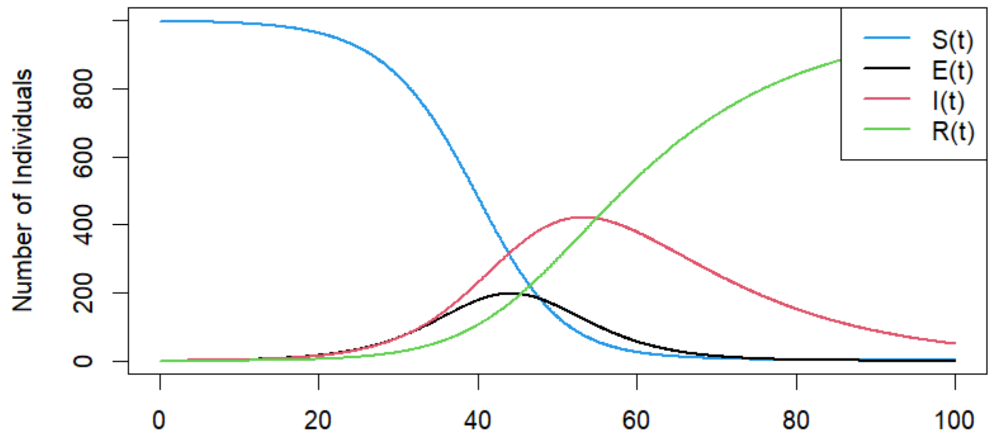
$where\ N = S + E + I + R\ is\ the\ total\ population$

| Parameter | Definition | True Value |
|---|---|---|
| $\beta$ | Infectious Rate | 0.400 |
| $\sigma$ | Incubation Rate | 0.200 |
| $\gamma$ | Recovery Rate | 0.059 |

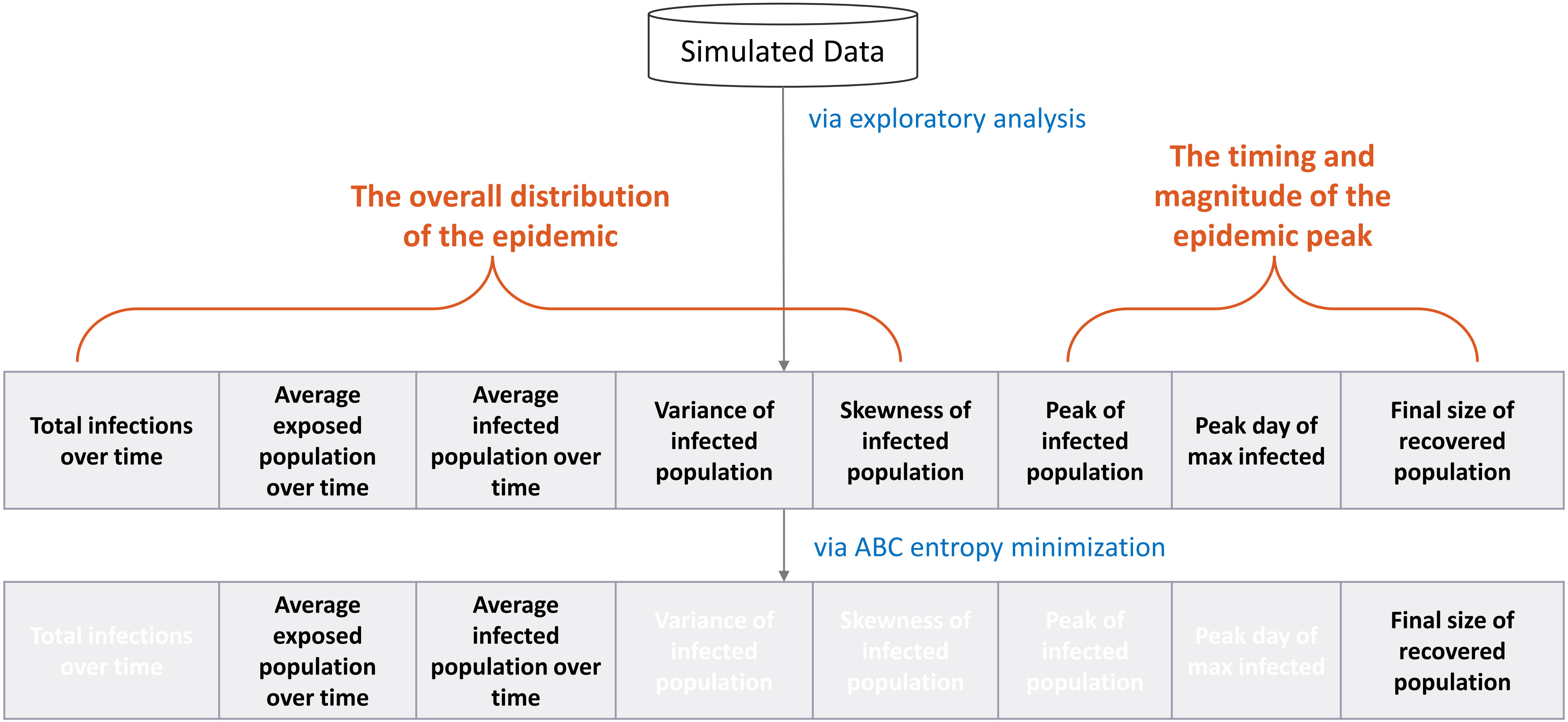
$Initial\ condition: S = 999, E = 0, I = 1, R = 0$

$Time: 100\ days$

**Plot of each of the variables changes through time**

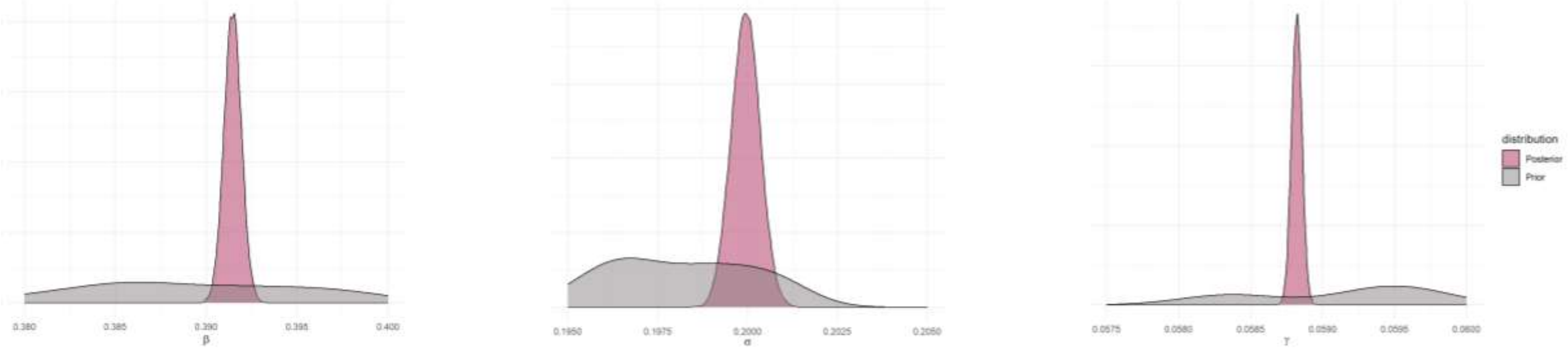


# 4.1 Simulation – Summary Statistics Selection using ABC

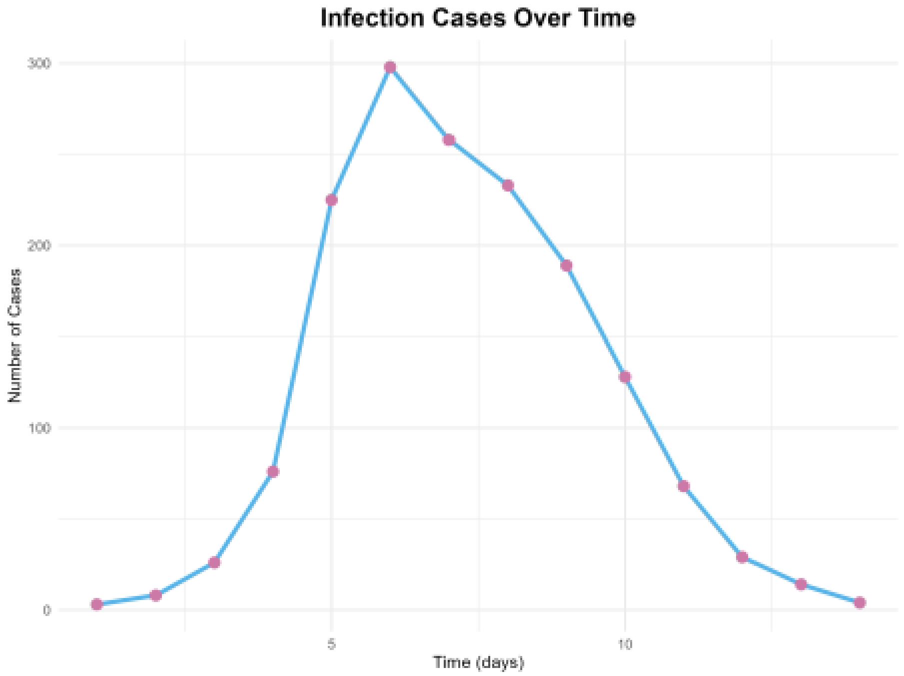


# 4.1 Simulation – Convergence Diagnostics

| Parameter | True Value | Mean | 2.5% | 97.5% | ESS | Rhat |
|---|---|---|---|---|---|---|
| $\beta$ | 0.4000 | 0.4000 | 0.3968 | 0.4034 | 1650 | 1.0026 |
| $\sigma$ | 0.2000 | 0.1999 | 0.1975 | 0.2026 | 1594 | 1.0032 |
| $\gamma$ | 0.0588 | 0.0588 | 0.0586 | 0.0590 | 3107 | 1.0012 |



# 4.2 Application – Data

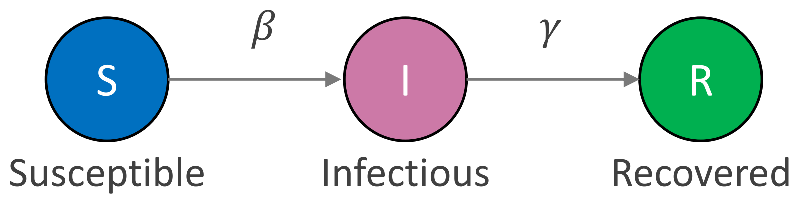



**About the dataset:**

- 1978 influenza A outbreak at a British boarding school.
- During the outbreak, 512 out of 763 students became ill.
- The illness spread between January 22nd and February 4th.
- The data for this outbreak is publicly available in the R package *outbreak*.

# 4.2 Application – Stochastic SIR Model (Chatzilena et al., 2019)

**Deterministic Part**




$$\frac{dS}{dt} = -\beta \frac{I(t)}{N} S(t)$$

$$\frac{dI}{dt} = \beta \frac{I(t)}{N} S(t) - \gamma I(t)$$

$$\frac{dR}{dt} = \gamma I(t)$$

**Stochastic Part**

$$Y_t \sim Poisson(\lambda_t)$$

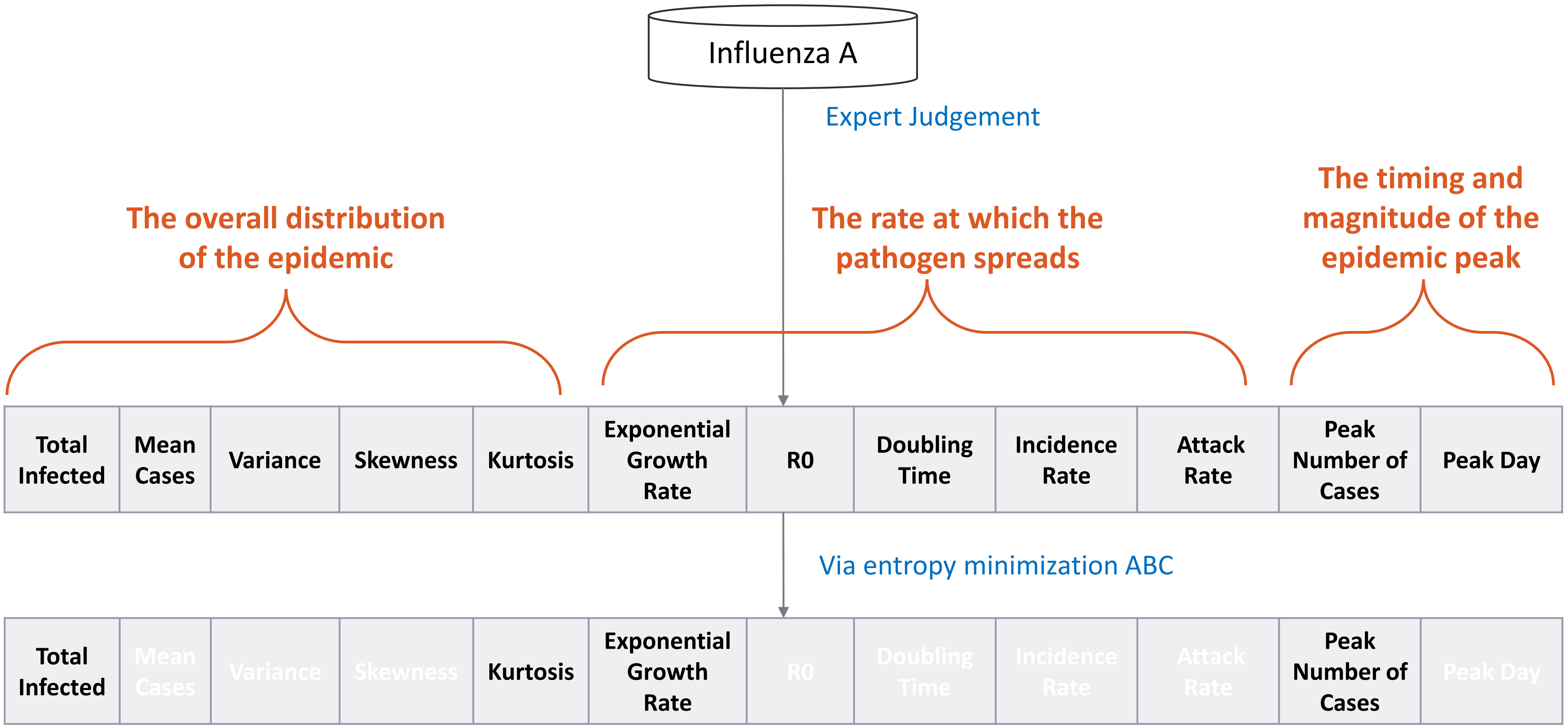
$$\lambda_t = exp(\kappa_t)$$

$$d\kappa_s = \phi(\mu_t - \kappa_s)dt + \sigma dB_s$$

$$\mu_t = log(\int (\beta \frac{I(s)}{N} S(s) - \gamma I(s))ds)$$

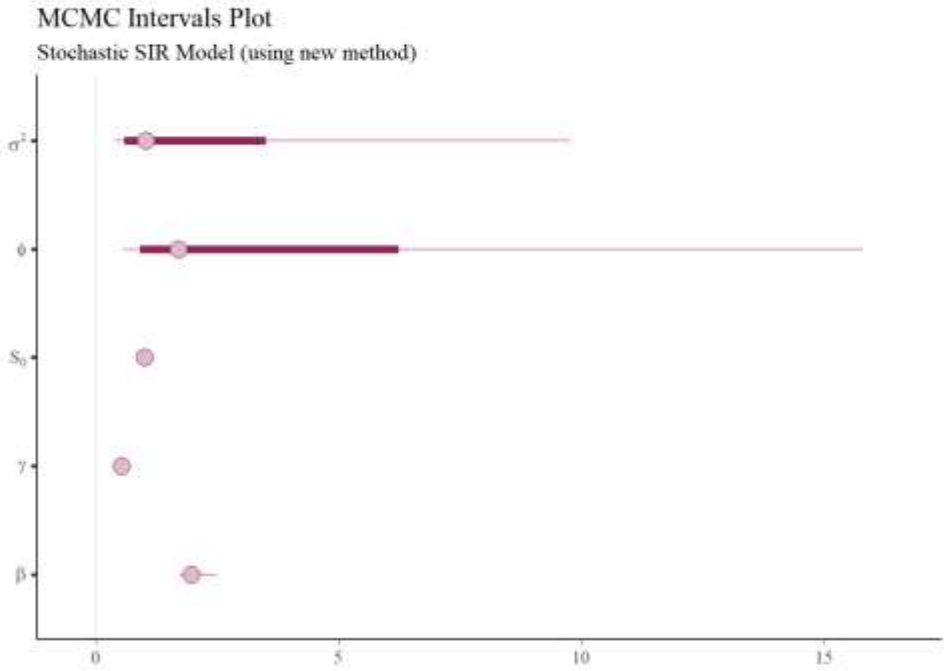
$$\kappa_{t+1} | \kappa_t \sim N(\mu_t + (\kappa_t - \mu_t)e^{-\phi}, \frac{\sigma^2}{2\phi}(1 - e^{-2\phi}))$$

# 4.2 Application – Summary Statistics Selection using ABC

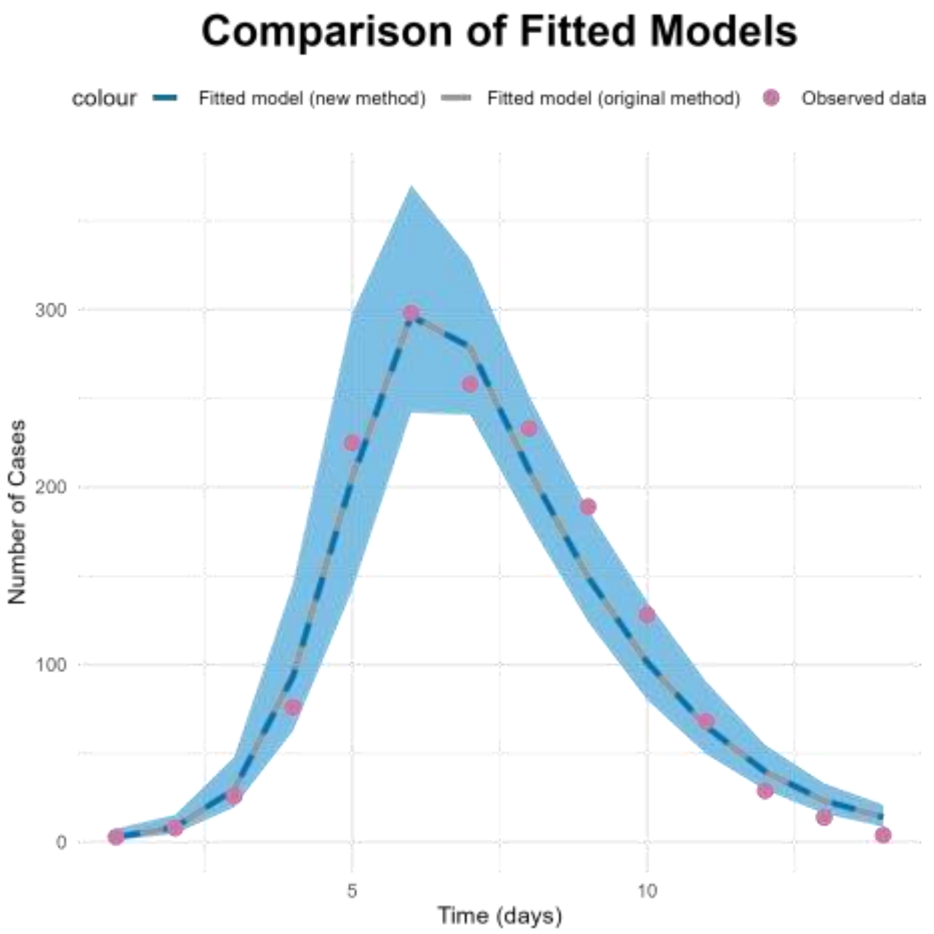


# 4.2 Application – Convergence Diagnostics

| Parameter | Definition | Prior Distribution |
|---|---|---|
| $\beta$ | Transmission rate | $Lognormal(0,1)$ |
| $\gamma$ | Recovery rate | $\Gamma(0.004, 0.002)$ |
| $s(0)$ | Initial proportion of susceptible individuals | $Beta(0.5, 0.5)$ |
| $\phi$ | The speed of inversion | $HalfNormal(0, 100)$ |
| $\sigma^2$ | The instantaneous diffusion term | $\Gamma^{-1}(0.1, 0.1)$ |



# 4.2 Application – Methods Comparison




| Method | Parameter | Mean | SD | 2.5% | 97.5% | ESS | Rhat |
|---|---|---|---|---|---|---|---|
| New | $\beta$ | 2.0161 | 0.2639 | 1.6993 | 2.7484 | 1550 | 1.0020 |
| | $\gamma$ | 0.5333 | 0.0534 | 0.4388 | 0.6561 | 2740 | 1.0002 |
| | $S_0$ | 0.9991 | 0.0006 | 0.9977 | 0.9998 | 1969 | 1.0017 |
| | $\sigma^2$ | 2.5427 | 3.2900 | 0.3400 | 12.0955 | 1484 | 1.0028 |
| | $\phi$ | 4.2612 | 5.1969 | 0.4621 | 19.0125 | 1382 | 1.0023 |
| | Time (min) | 0.5591 | | | | | |
| | Memory (MB) | 6.5140 | | | | | |
| Original | $\beta$ | 2.0210 | 0.3289 | 1.6995 | 2.6647 | 746 | 1.0058 |
| | $\gamma$ | 0.5317 | 0.0503 | 0.4431 | 0.6425 | 2404 | 0.9998 |
| | $S_0$ | 0.9991 | 0.0005 | 0.9978 | 0.9998 | 1809 | 1.0003 |
| | $\sigma^2$ | 2.5261 | 3.3611 | 0.3538 | 12.0061 | 1229 | 1.0010 |
| | $\phi$ | 4.2190 | 5.0453 | 0.4898 | 18.2884 | 1005 | 1.0013 |
| | Time (min) | 1.6304 | | | | | |
| | Memory (MB) | 5.4893 | | | | | |

# 5 Discussion - Summary

**Summary:**

- Introduced a **hybrid inference framework** for parameter estimation in epidemiological models with intractable likelihoods.
- Addressed challenges in Bayesian inference for compartmental models, by integrating
  - **ABC-based entropy minimization** for summary statistics selection
  - **BSL** for flexible likelihood approximation
  - **HMC** for posterior sampling

# 5 Discussion – Limitation and Further Research

- Current numerical studies only focused on **compartmental models**, can extend to **more complex models** (e.g., agent-based).
- **Gaussian approximation** might be restrictive for extreme cases, can explore **alternative assumptions** to improve **robustness.**
- Apply to **large-scale, real-world datasets** to test scalability and performance.
- **Theoretical investigation** on method under varying degrees of model complexity and data quality.

# 6 Reference


Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, *162*(4), 2025-2035.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:1701.02434.

Bjørnstad, O. N., Shea, K., Krzywinski, M., & Altman, N. (2020). The SEIRS model for infectious disease dynamics. Nature methods, 17(6), 557-559.

Brauer, F. (2008). Compartmental models in epidemiology. *Mathematical epidemiology*, 19-79.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*.

Fearnhead, P., & Prangle, D. (2012). Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. Journal of the Royal Statistical Society Series B: Statistical Methodology, 74(3), 419-474.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). Bayesian data analysis. Chapman and Hall/CRC.

Li, X., Swallow, B., & Chadwick, F. J. (2024). A Novel Approximate Bayesian Inference Method for Compartmental Models in Epidemiology using Stan. *arXiv preprint arXiv:2408.03415*.

Lopes, J. S., & Beaumont, M. A. (2010). ABC: a useful Bayesian tool for the analysis of population data. Infection, Genetics and Evolution, 10(6), 825-832.

Minter, A., & Retkute, R. (2019). Approximate Bayesian Computation for infectious disease modelling. Epidemics, 29, 100368.

Nunes, M. A., & Prangle, D. (2015). abctools: an R package for tuning approximate Bayesian computation analyses. The R Journal.

Shearer, F. M., Edwards, L., Kirk, M., Eales, O., Golding, N., Hassall, J., ... & McCaw, J. M. (2024). Opportunities to strengthen respiratory virus surveillance systems in Australia: lessons learned from the COVID-19 response. Communicable Diseases Intelligence, 48.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., & Stumpf, M. P. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. Journal of the Royal Society Interface, 6(31), 187-202.

Wood, S. N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. Nature, 466(7310), 1102-1104.

# Thank You

Xiahui Li
xl94@st-andrews.ac.uk
University of St Andrews

Dr Fergus Chadwick
fergusjchadwick@st-andrews.ac.uk
University of St Andrews

Dr Ben Swallow
bts3@st-andrews.ac.uk
University of St Andrews