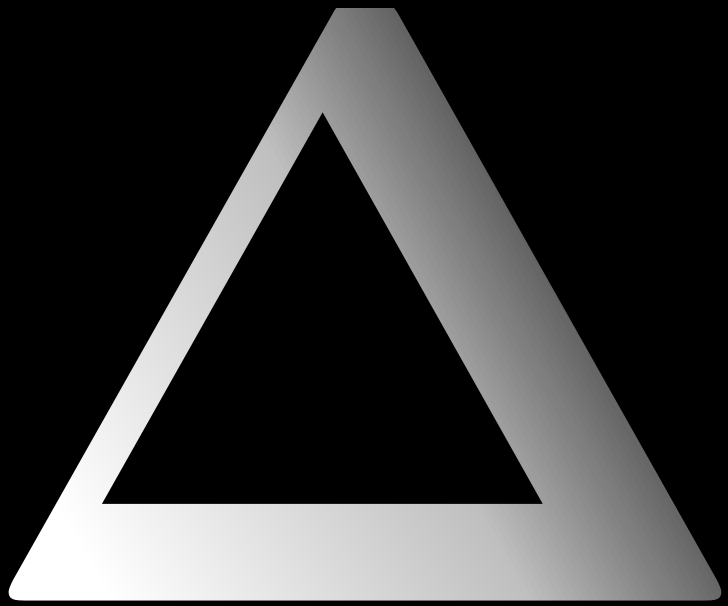
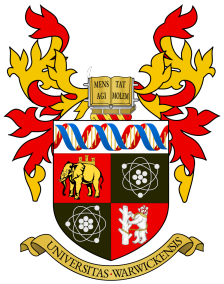




EPSRC

Engineering and Physical Sciences Research Council

Epidemics & Errors: Finding the Silver Lining

Zak Ogi-Gittins

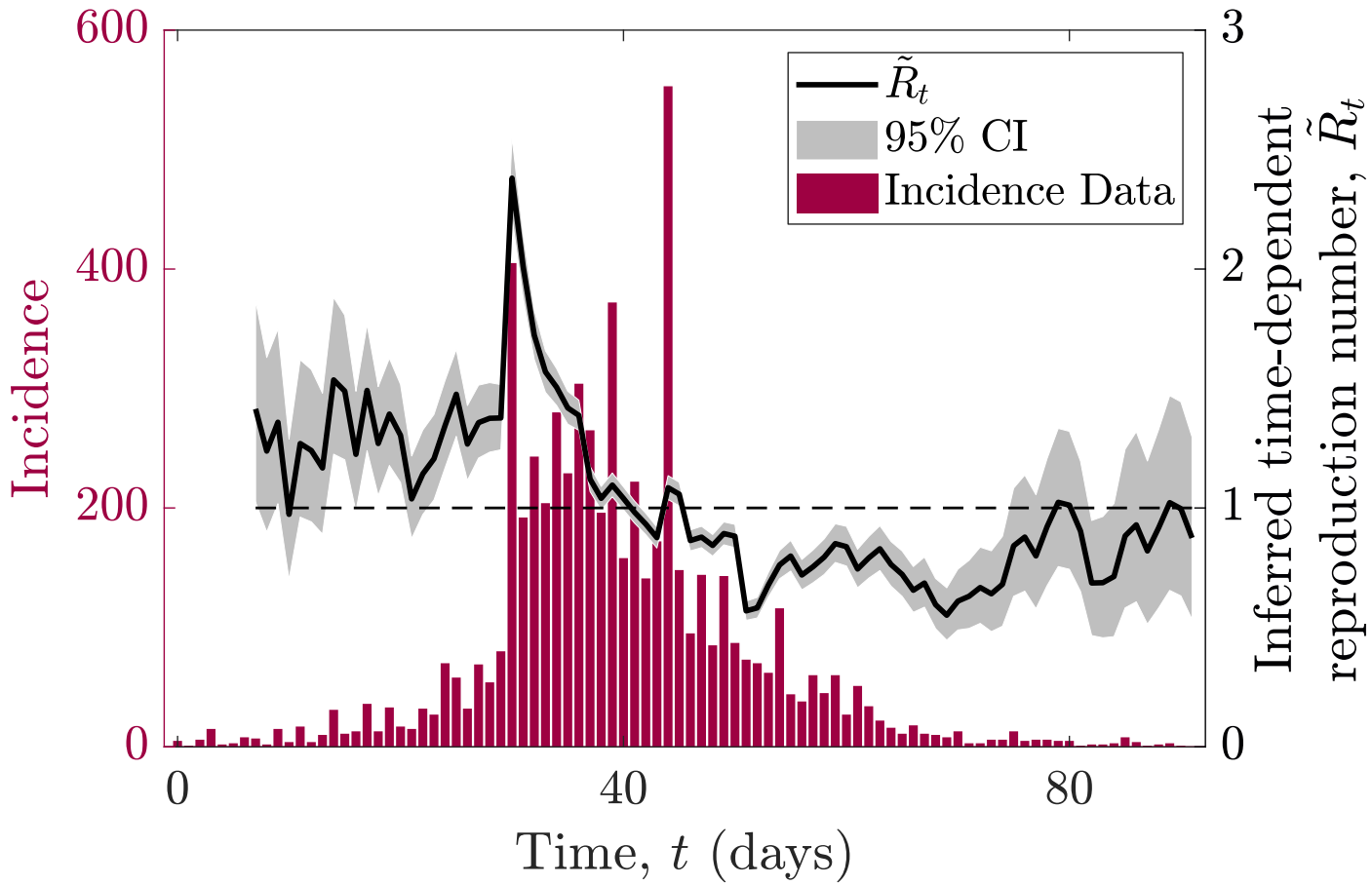
University of Warwick

zak.gittins@warwick.ac.uk

September 27, 2021

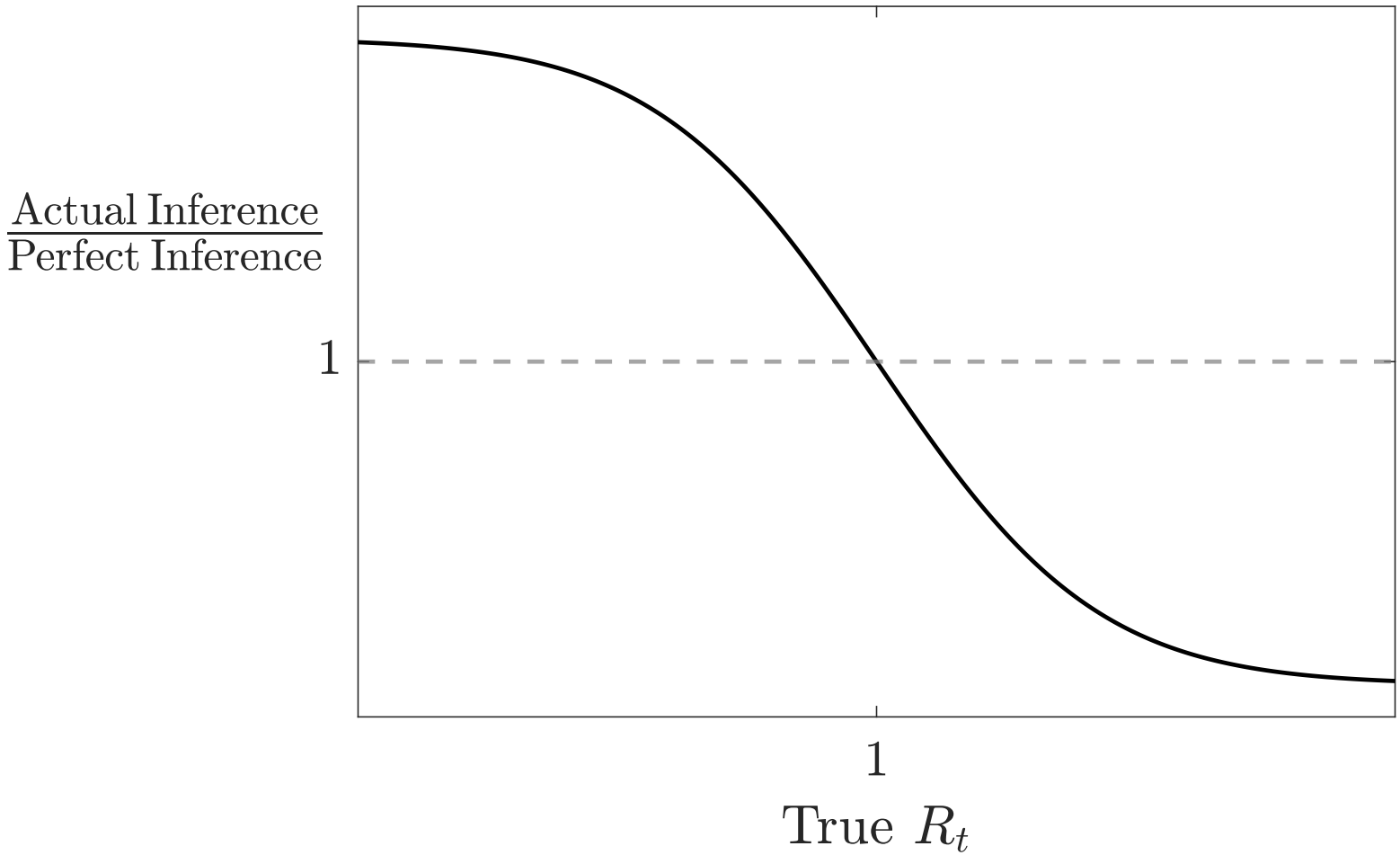
Imperfect data collection may introduce systematic bias into parameter inference

During epidemics, the reproduction number, R_t, is an important parameter to infer

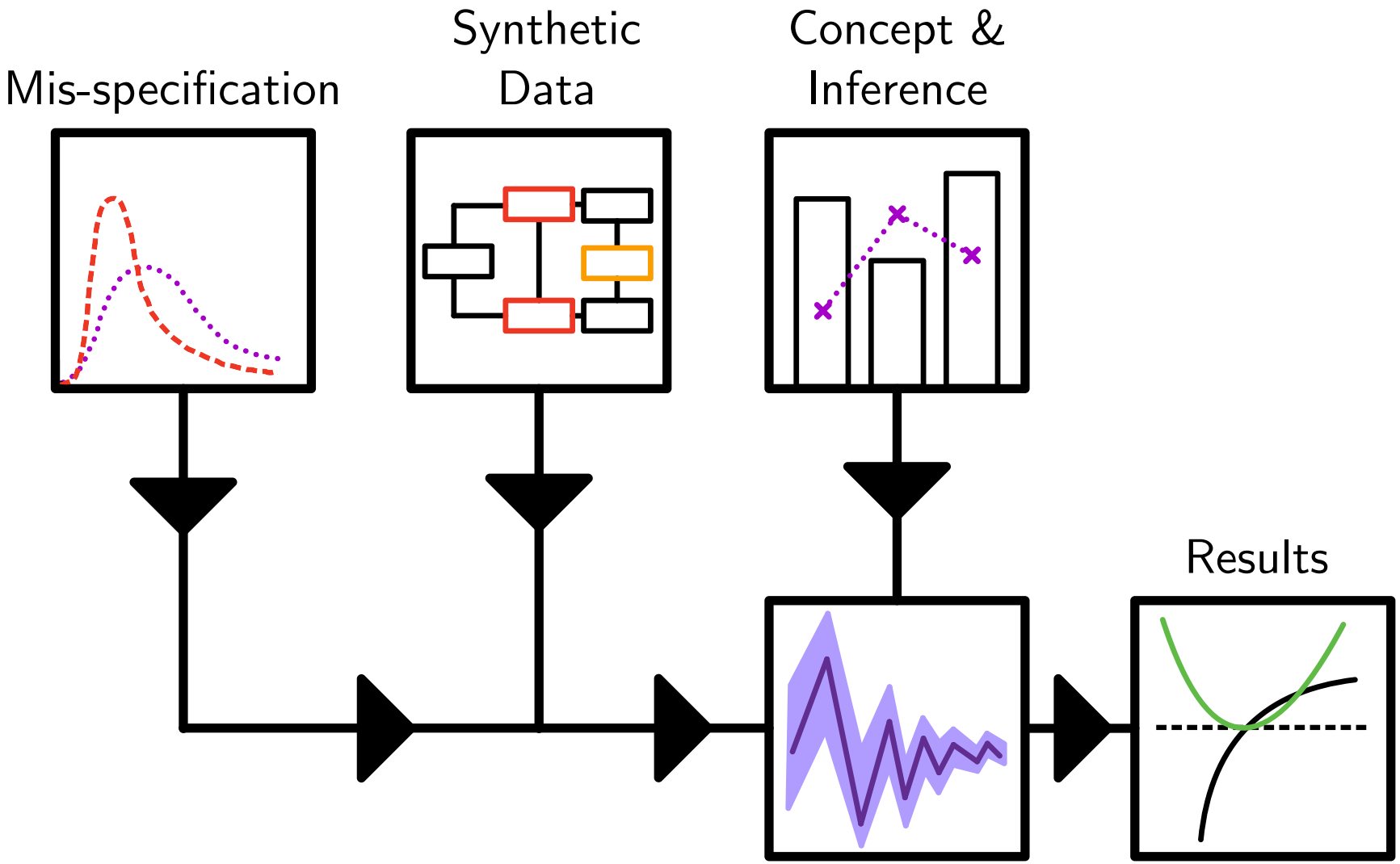


Data source: Cori et. al, AJE 2013

R_t inference may be improved with extra information about systematic inference bias



Talk Outline



Definitions

Definitions

- R_t: The average number of people someone infected at time t could expect to infect (should conditions remain unchanged).

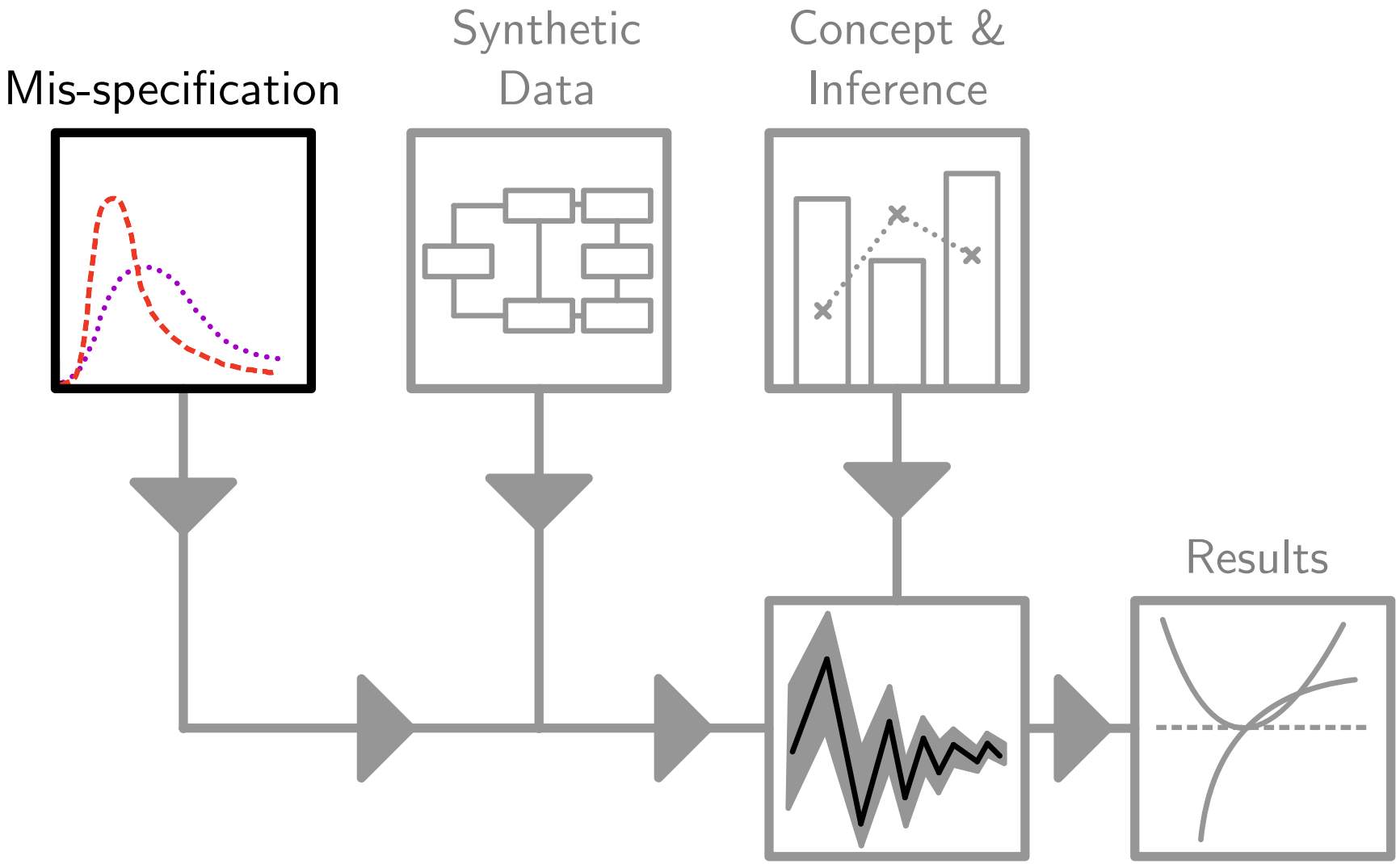
Definitions

- R_t: The average number of people someone infected at time t could expect to infect (should conditions remain unchanged).
- Generation interval (GI): The interval between the time when an individual is infected by an infector and the time when this infector was infected.

Definitions

- R_t: The average number of people someone infected at time t could expect to infect (should conditions remain unchanged).
- Generation interval (GI): The interval between the time when an individual is infected by an infector and the time when this infector was infected.
- Incidence: The number of individuals who develop a specific disease during a particular time period (we assume days).

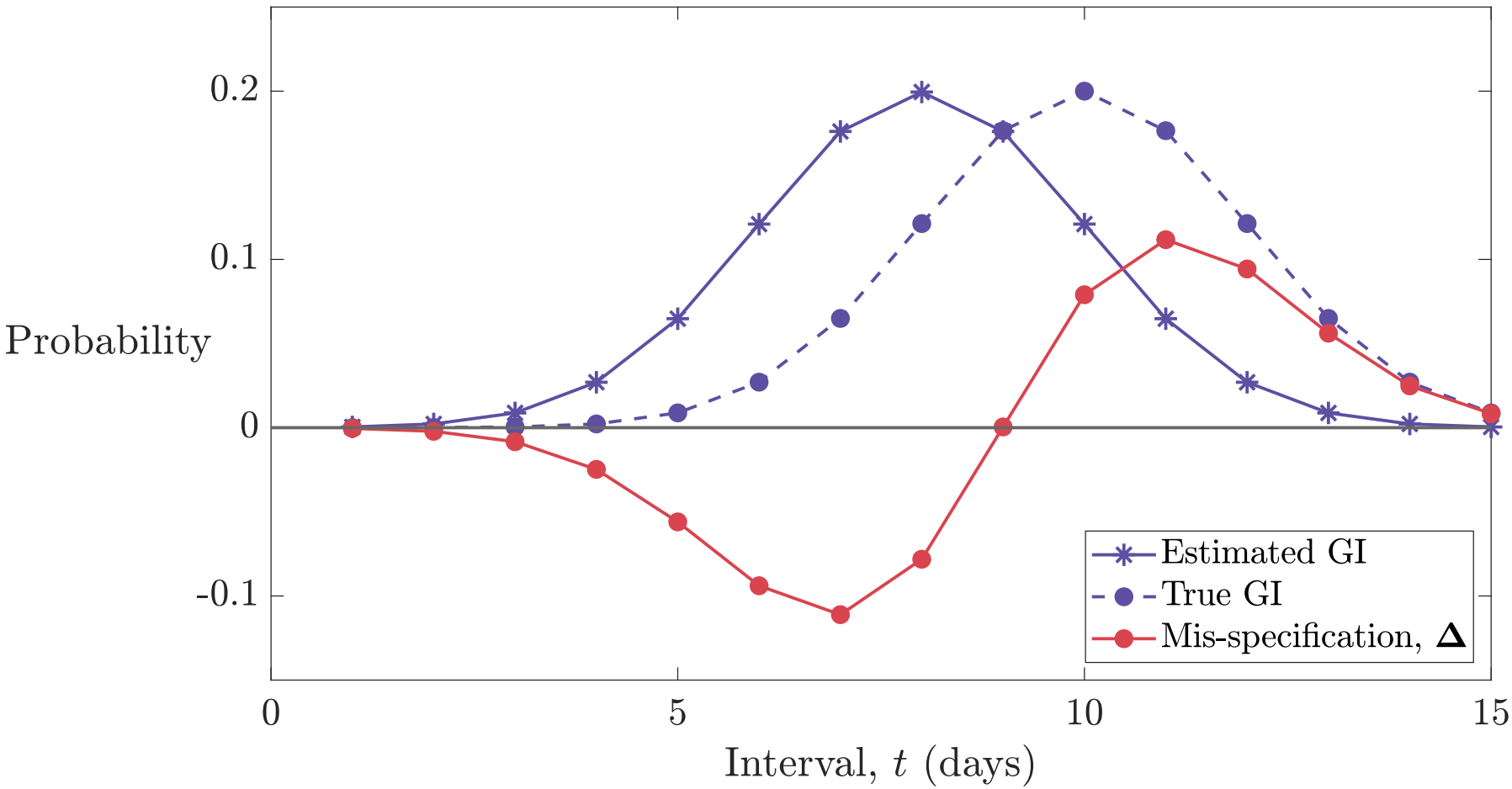
Mis-specification of generation intervals



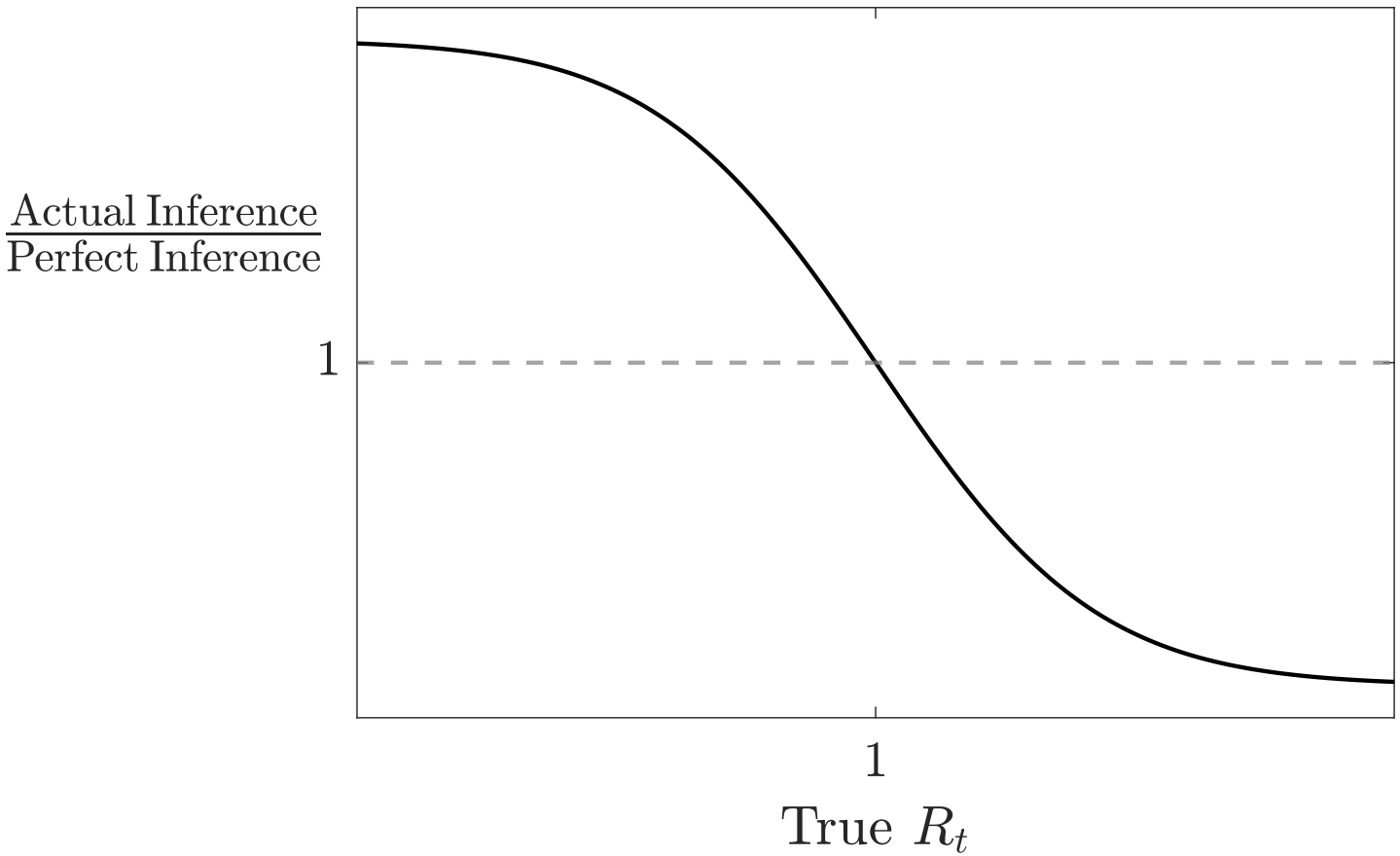
Generation intervals (GIs) can change during the course of an epidemic

One example cause: Public health measure e.g. enforced isolation if tested positive for disease.

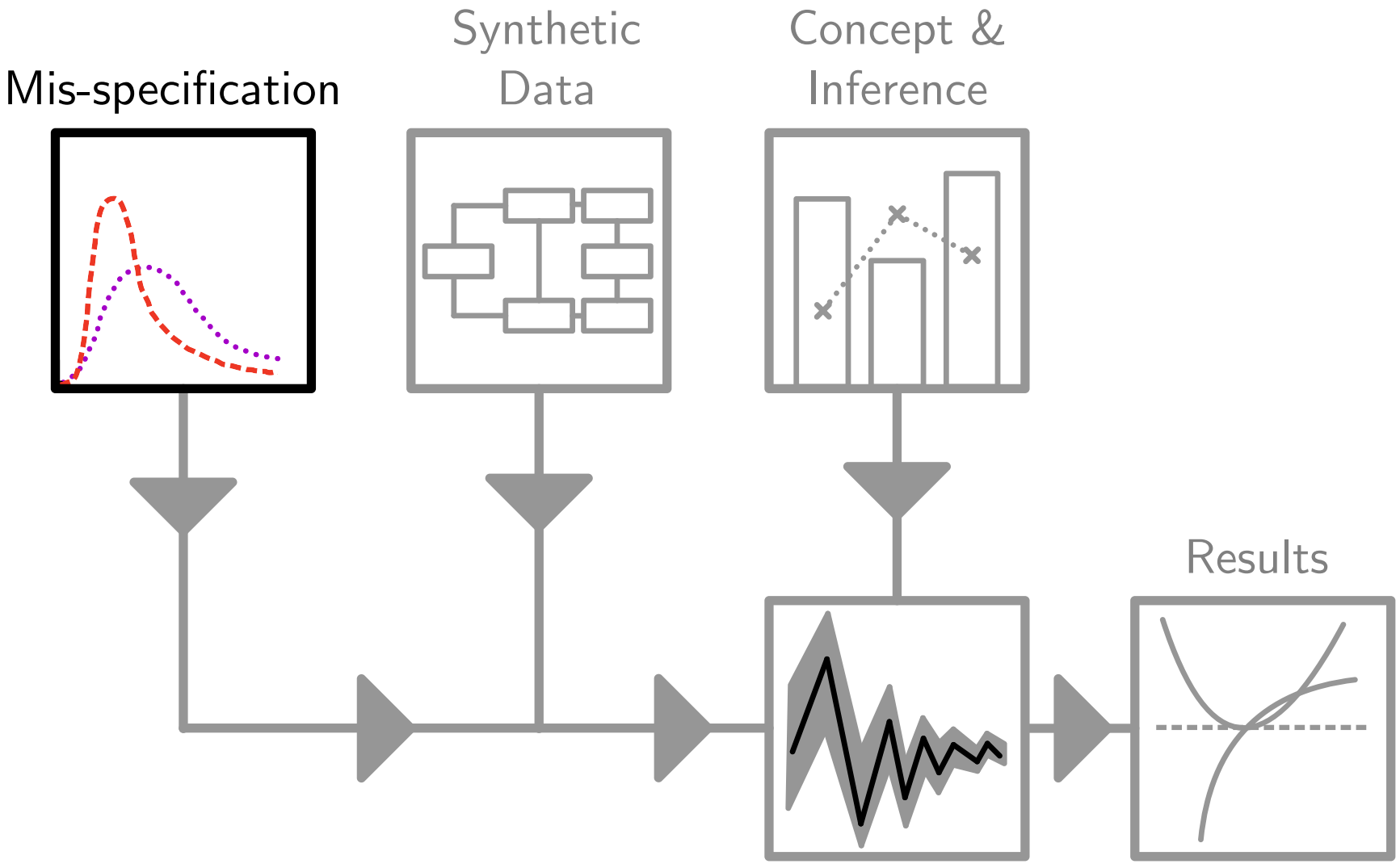
The mis-specification, Δ, is the difference between the recorded & true generation intervals



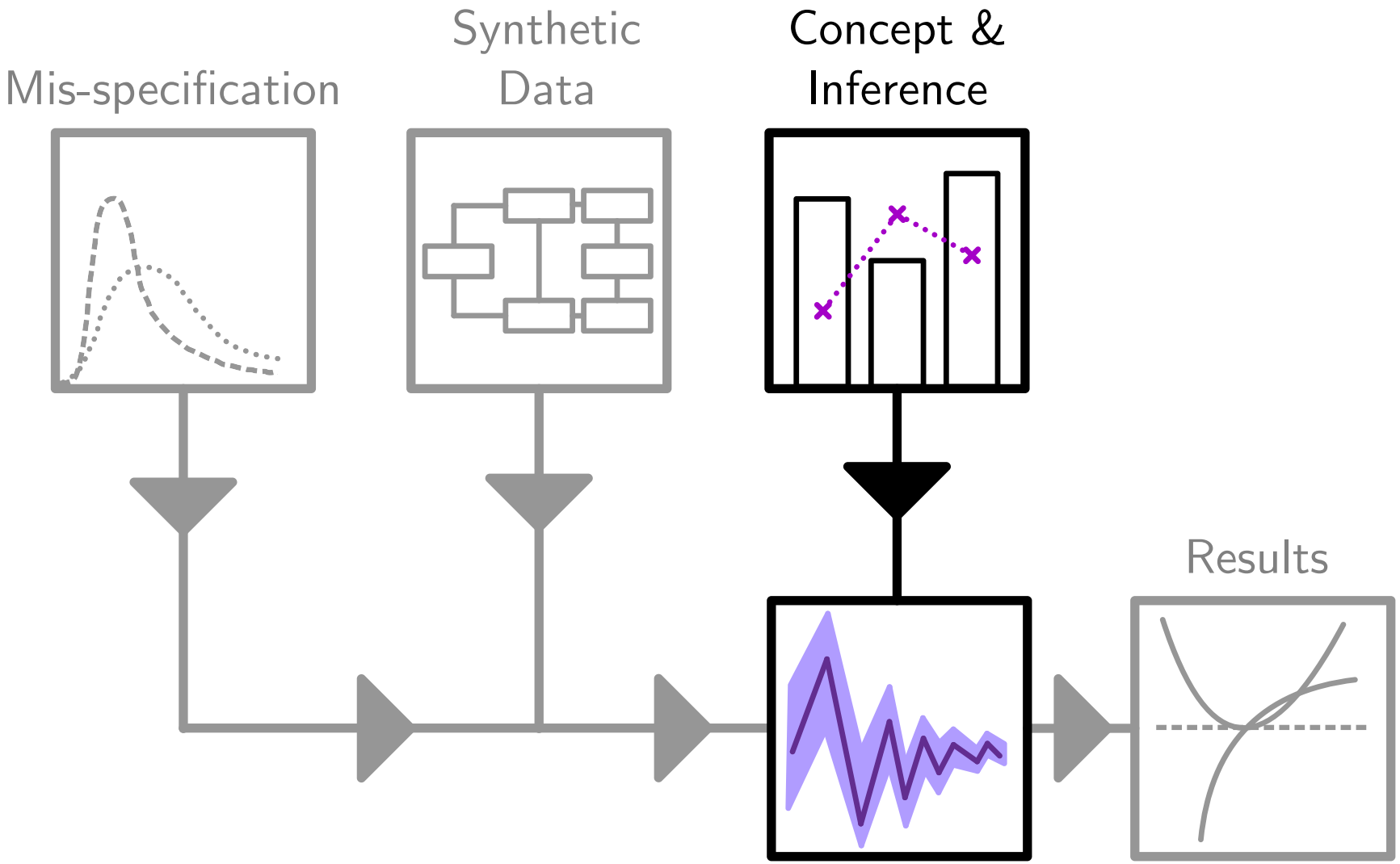
Certain mis-specification characteristics are significant in systematic R_t inference bias



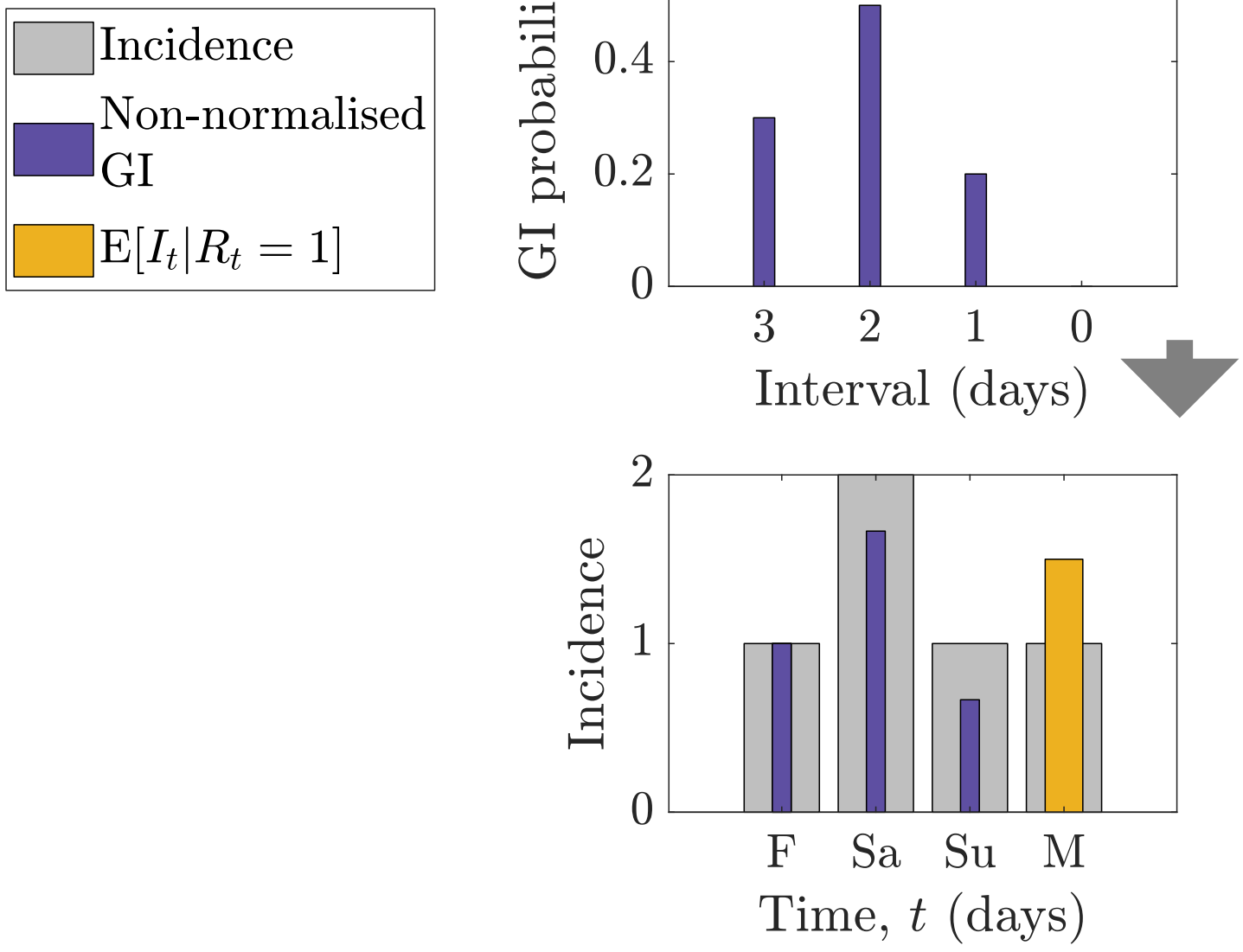
Mis-specification of generation intervals



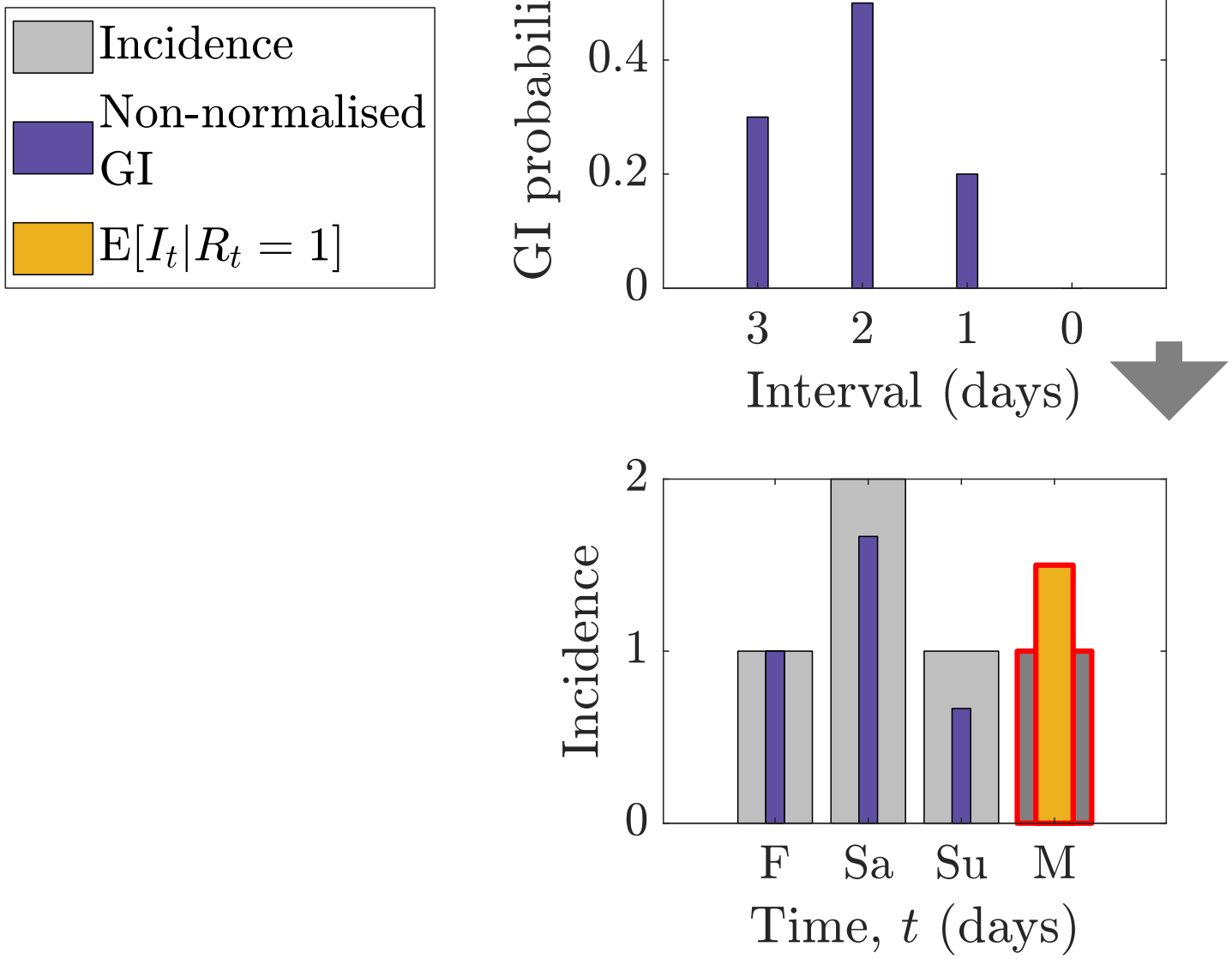
Concept for R_t inference



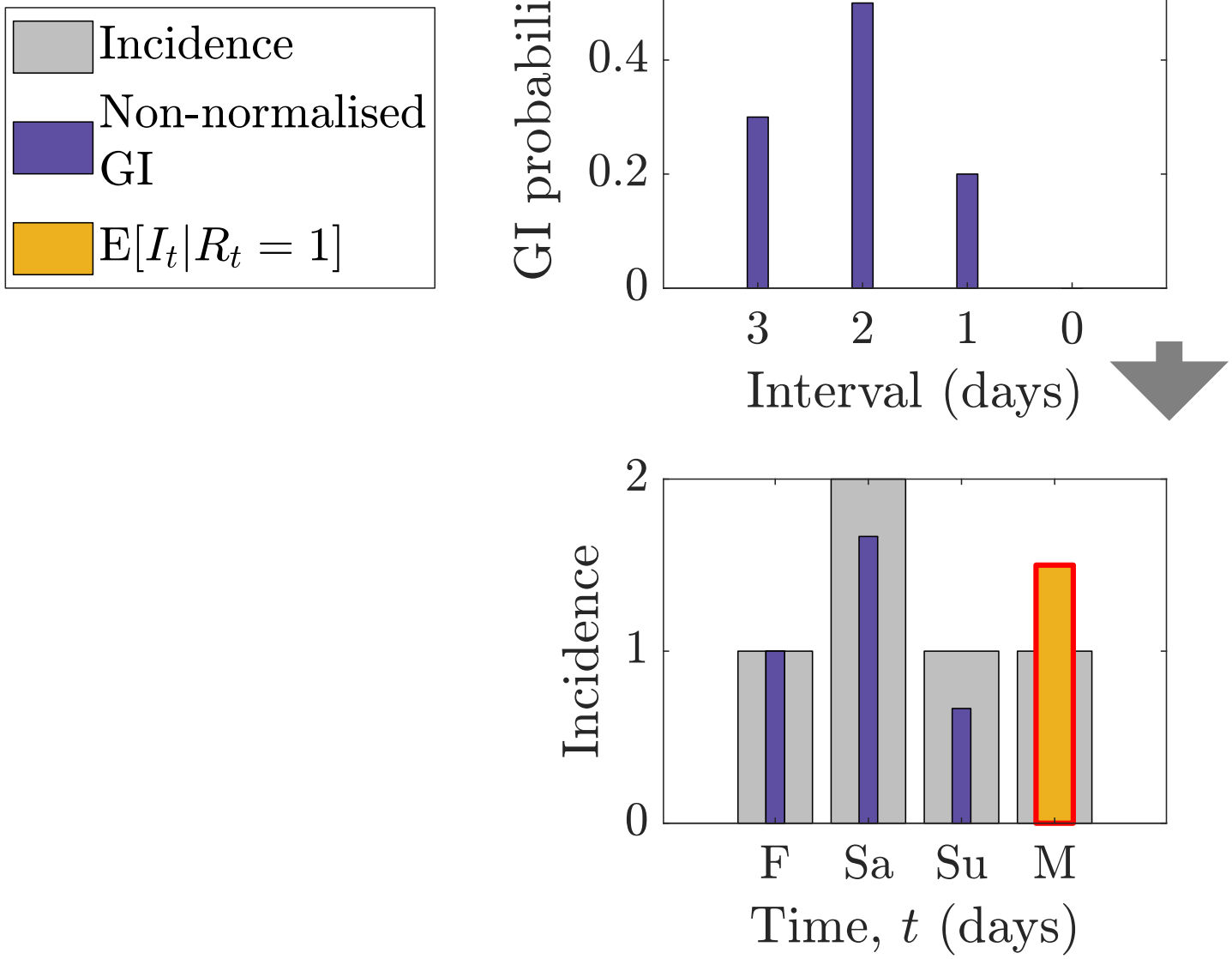
R_t can be estimated using estimated GIs & incidence data



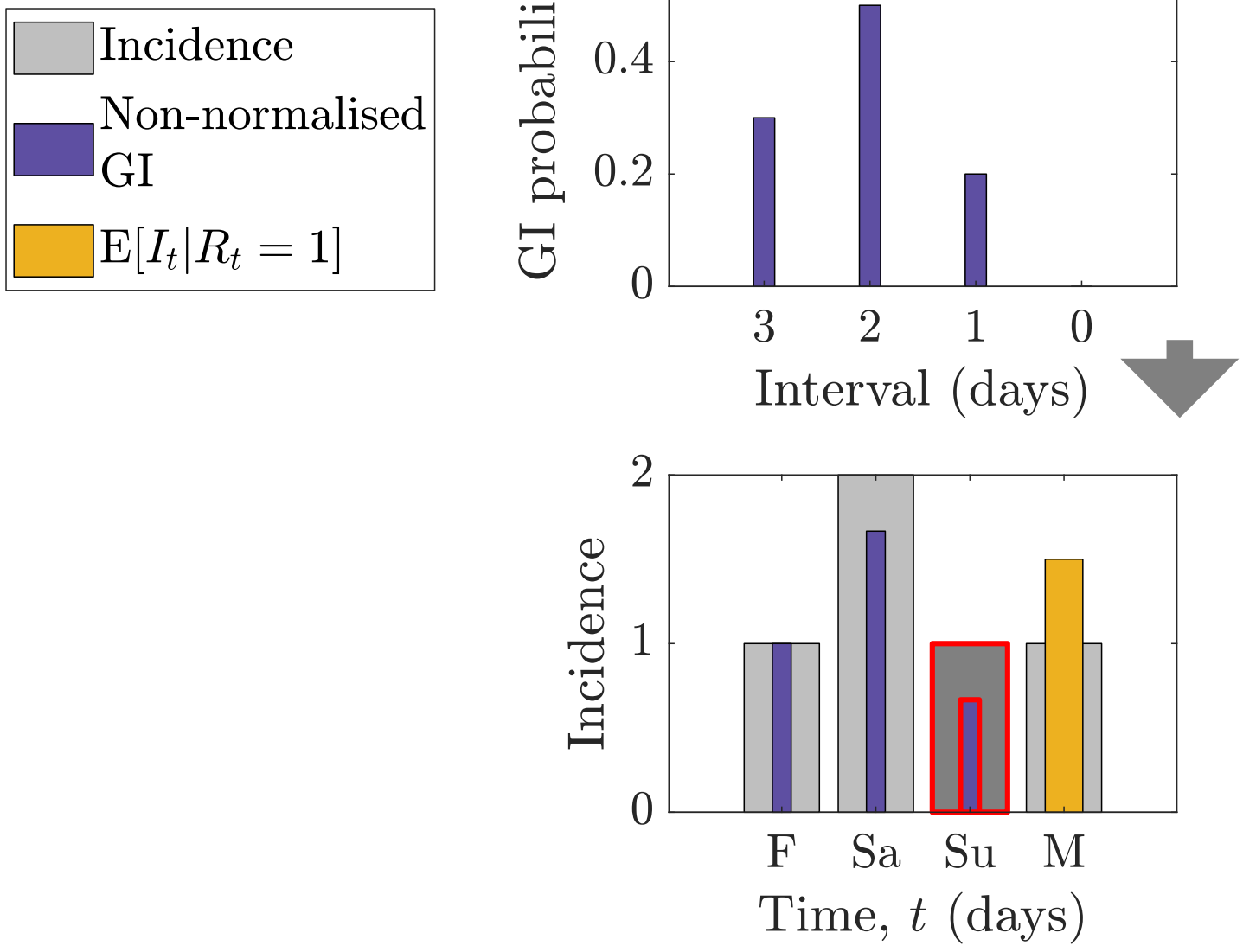
R_t can be estimated using estimated GIs & incidence data



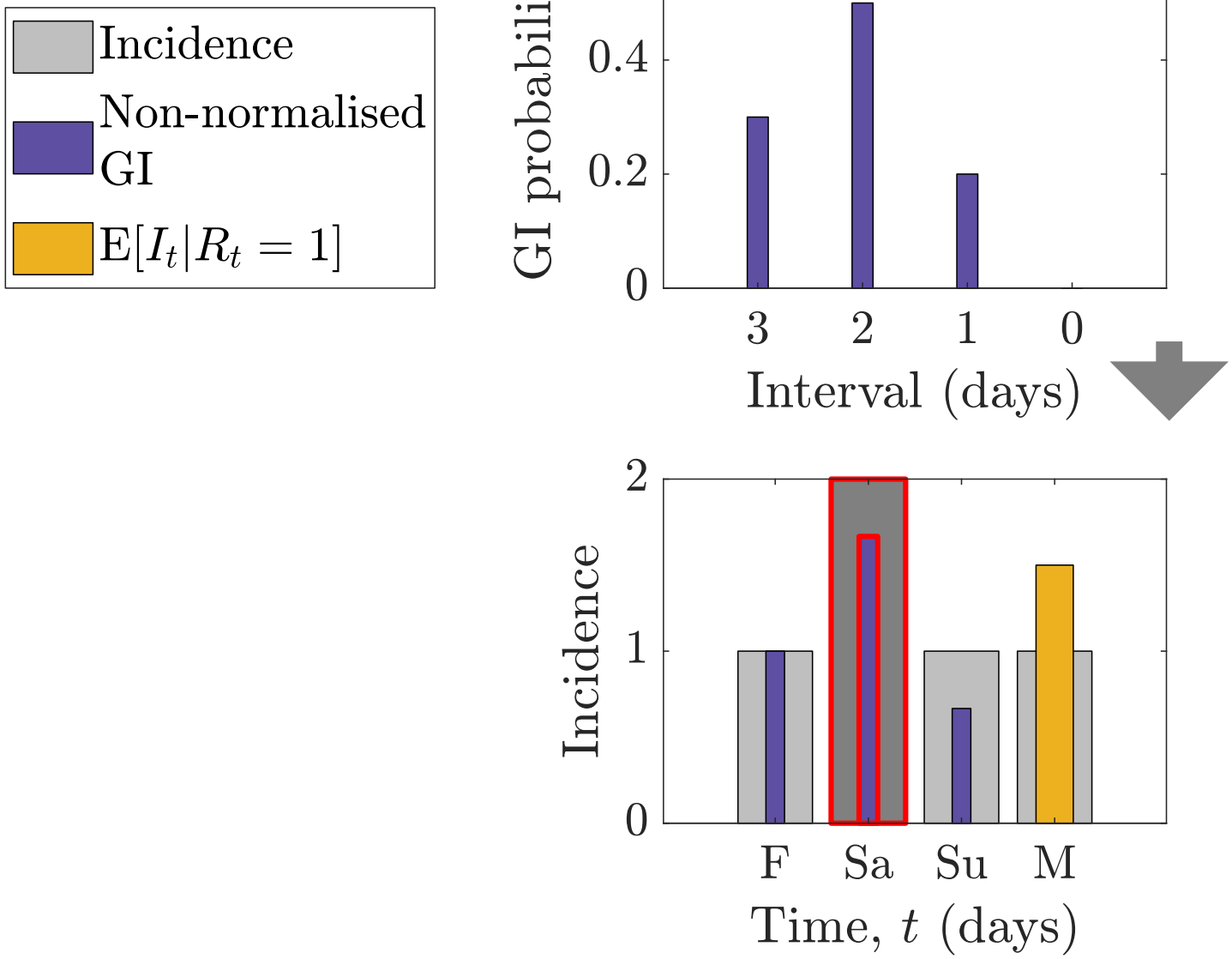
R_t can be estimated using estimated GIs & incidence data



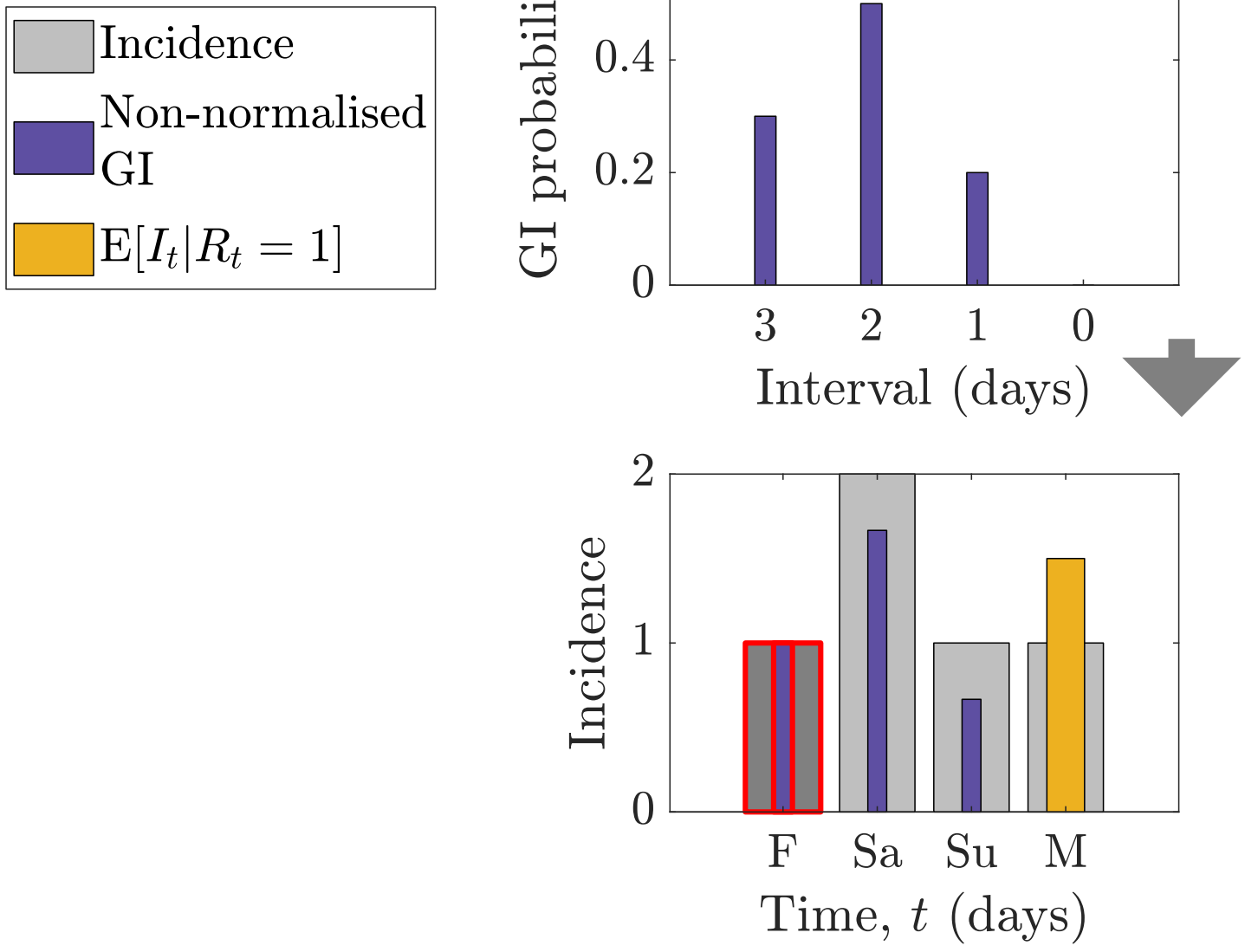
R_t can be estimated using estimated GIs & incidence data



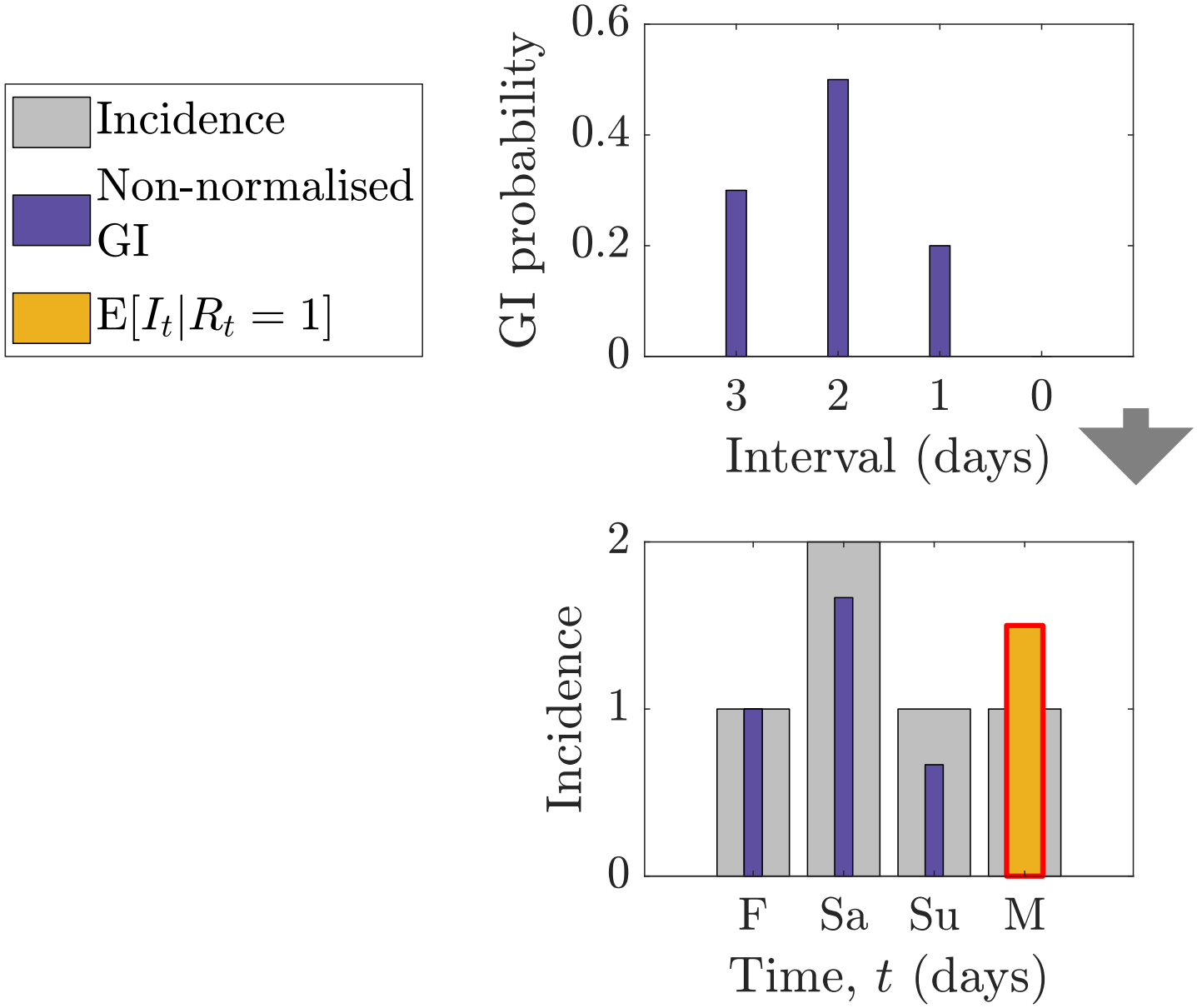
R_t can be estimated using estimated GIs & incidence data



R_t can be estimated using estimated GIs & incidence data



R_t can be estimated using estimated GIs & incidence data

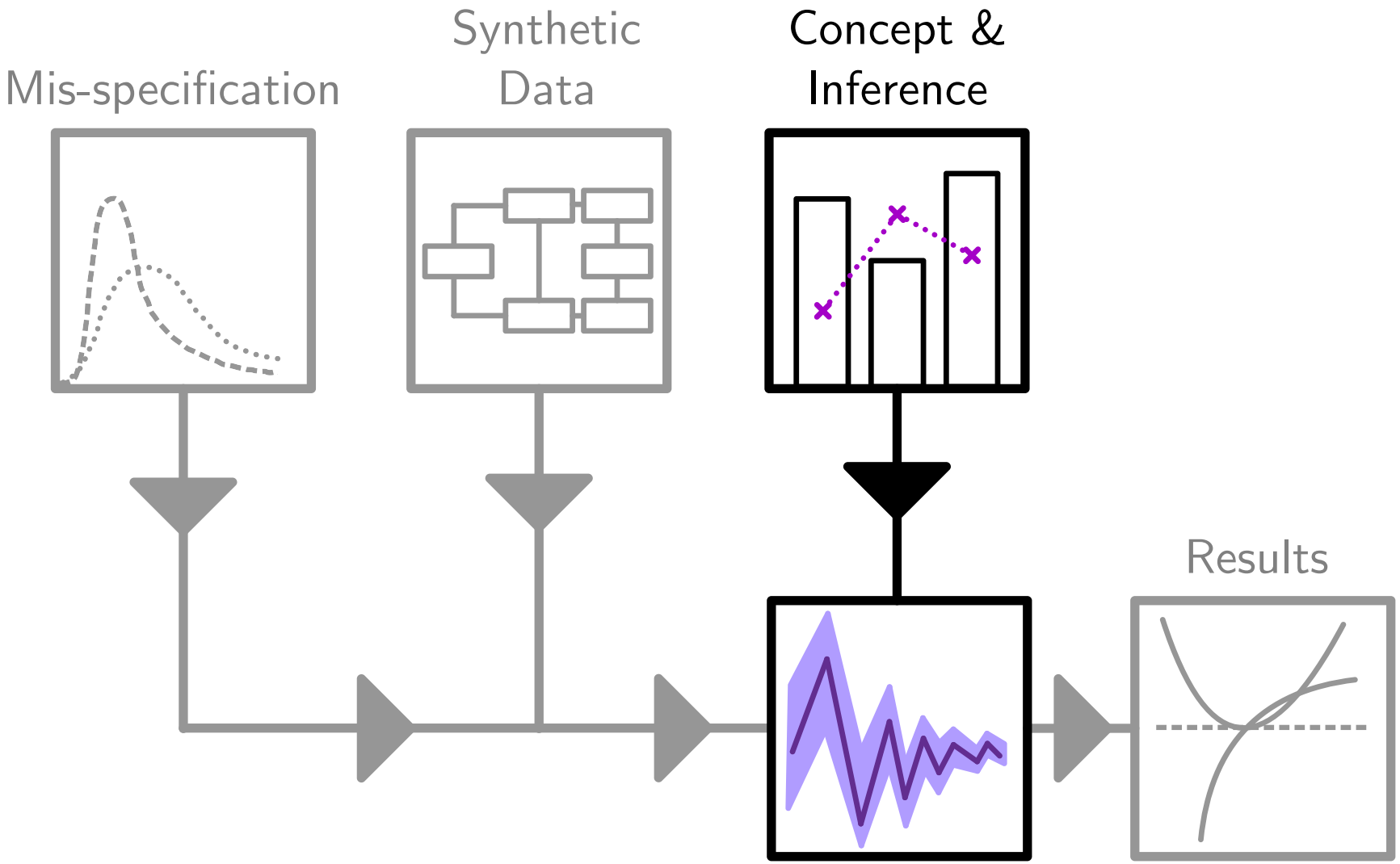


R_t can be estimated using estimated GIs & incidence data

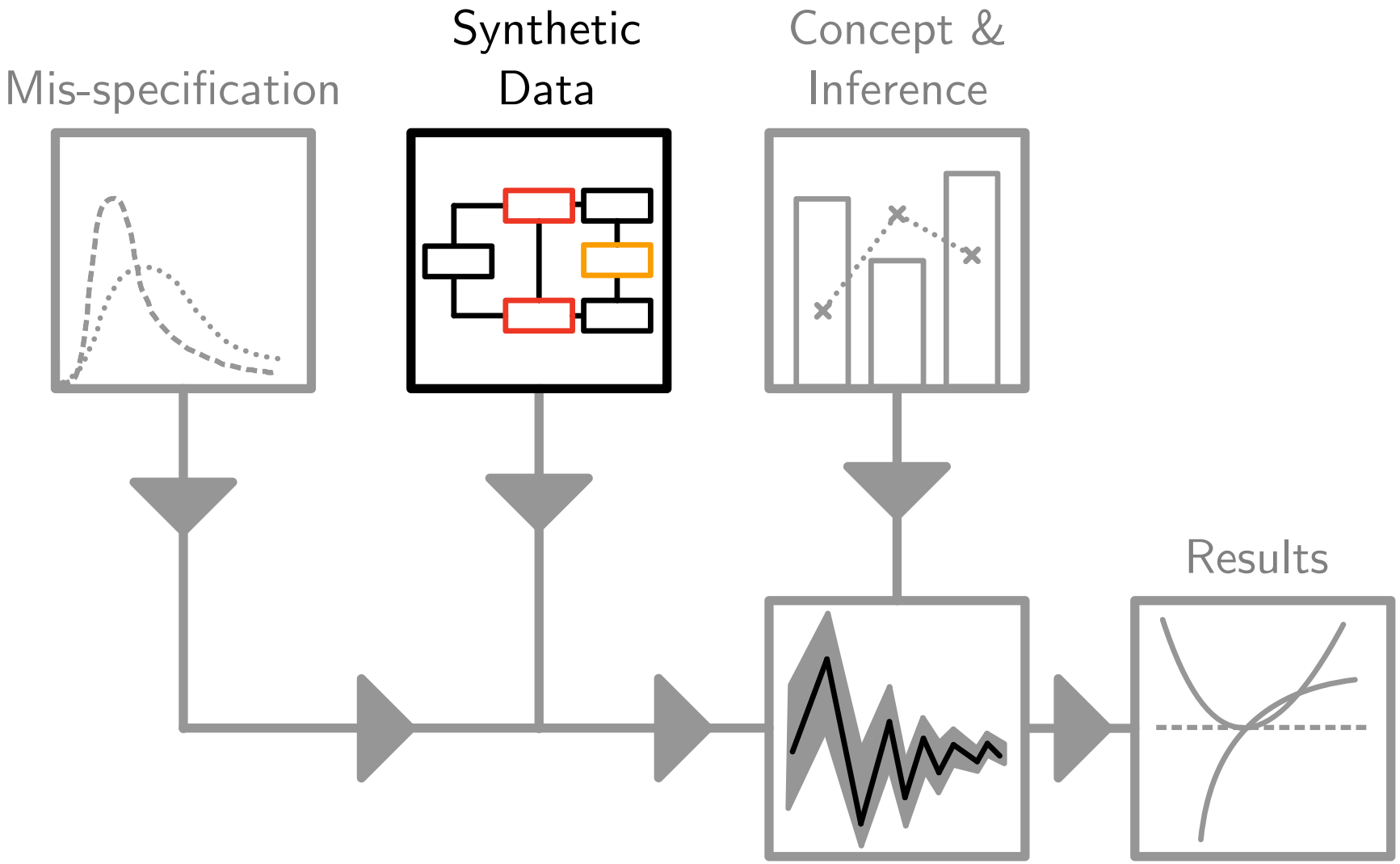
$$\begin{aligned}\mathbb{E}[I_{\text{Mon}}|R_{\text{Mon}}=1] &= \mathbb{P}[\text{GI}=1]\cdot I_{\text{Sun}} \\ &+ \mathbb{P}[\text{GI}=2]\cdot I_{\text{Sat}} \\ &+ \mathbb{P}[\text{GI}=3]\cdot I_{\text{Fri}}\end{aligned}$$

$$\tilde{R}_{\text{Mon}} = \frac{I_{\text{Mon}}}{\mathbb{E}[I_{\text{Mon}}|R_{\text{Mon}}=1]}$$

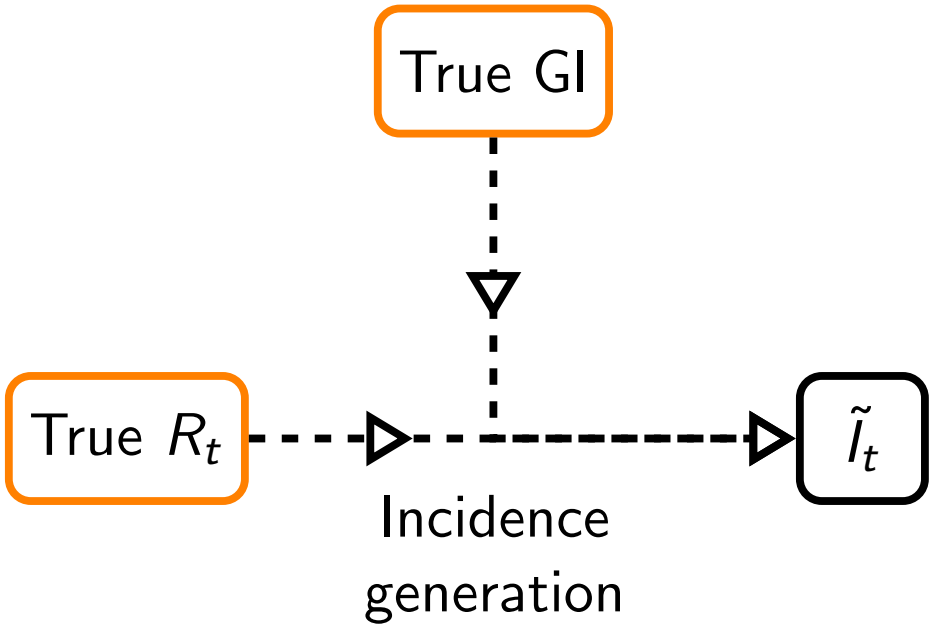
Concept for R_t inference



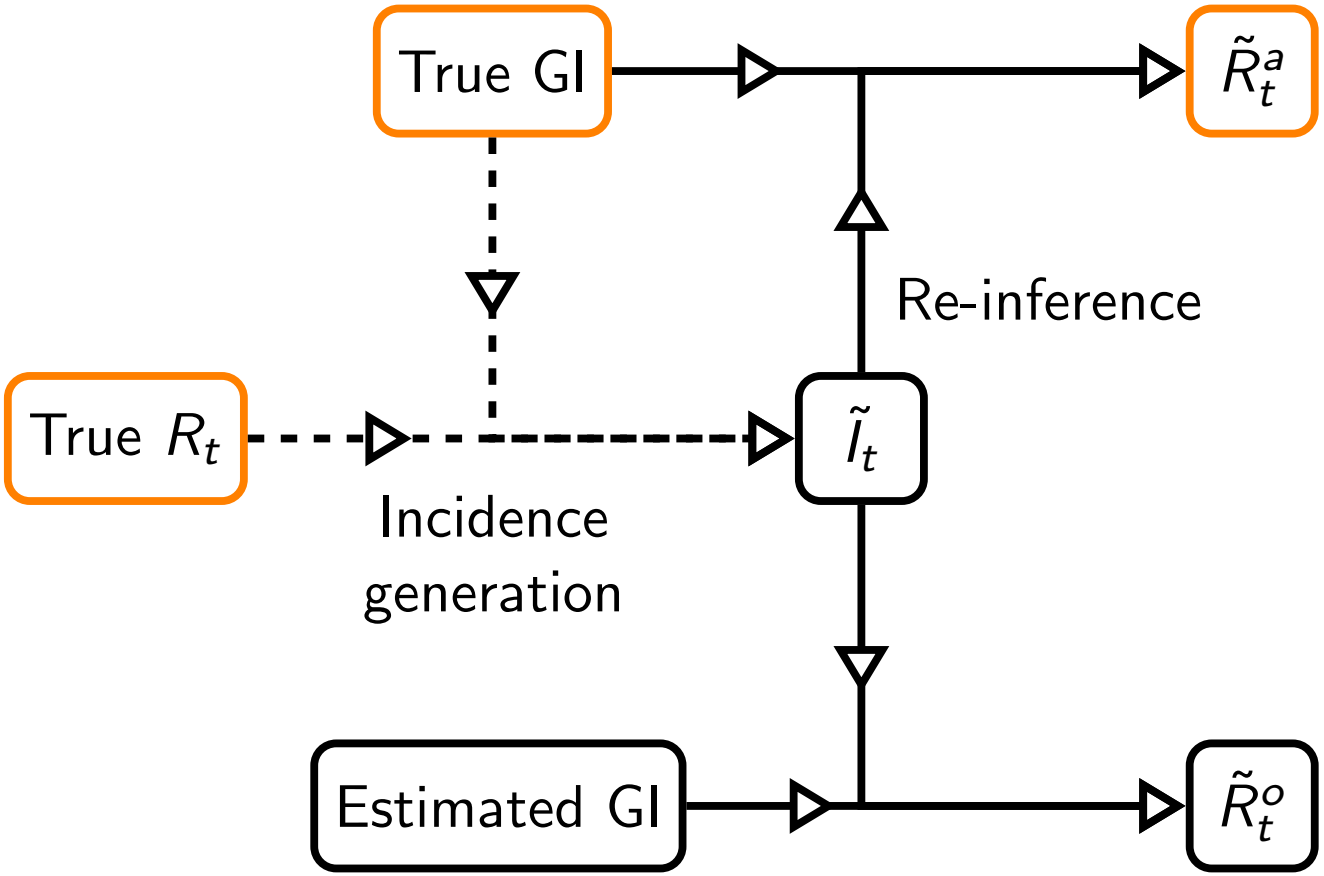
Comparing perfect & realistic inferences via synthetic data



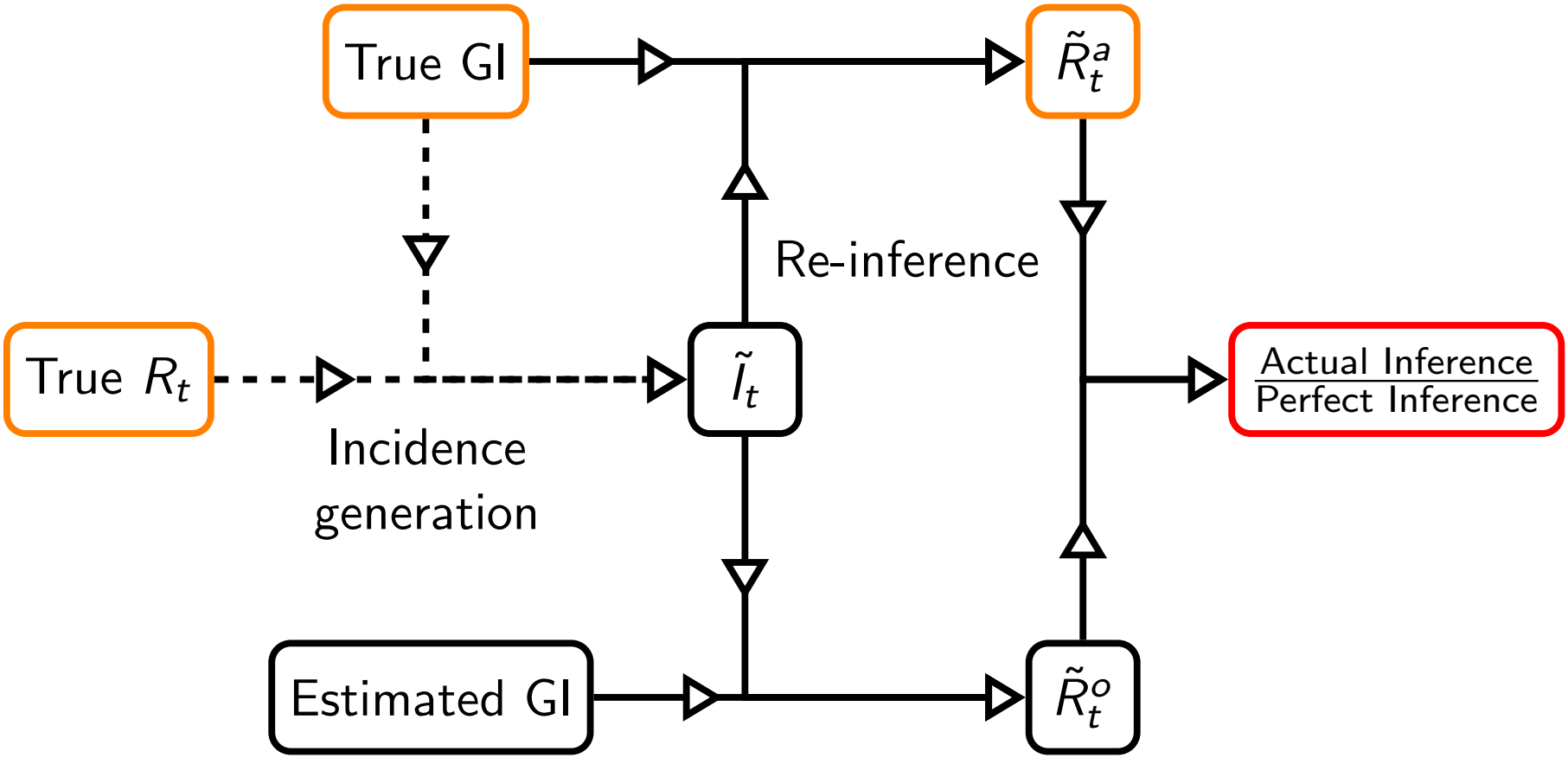
Synthetic outbreaks can be generated using the true R_t and the true GI



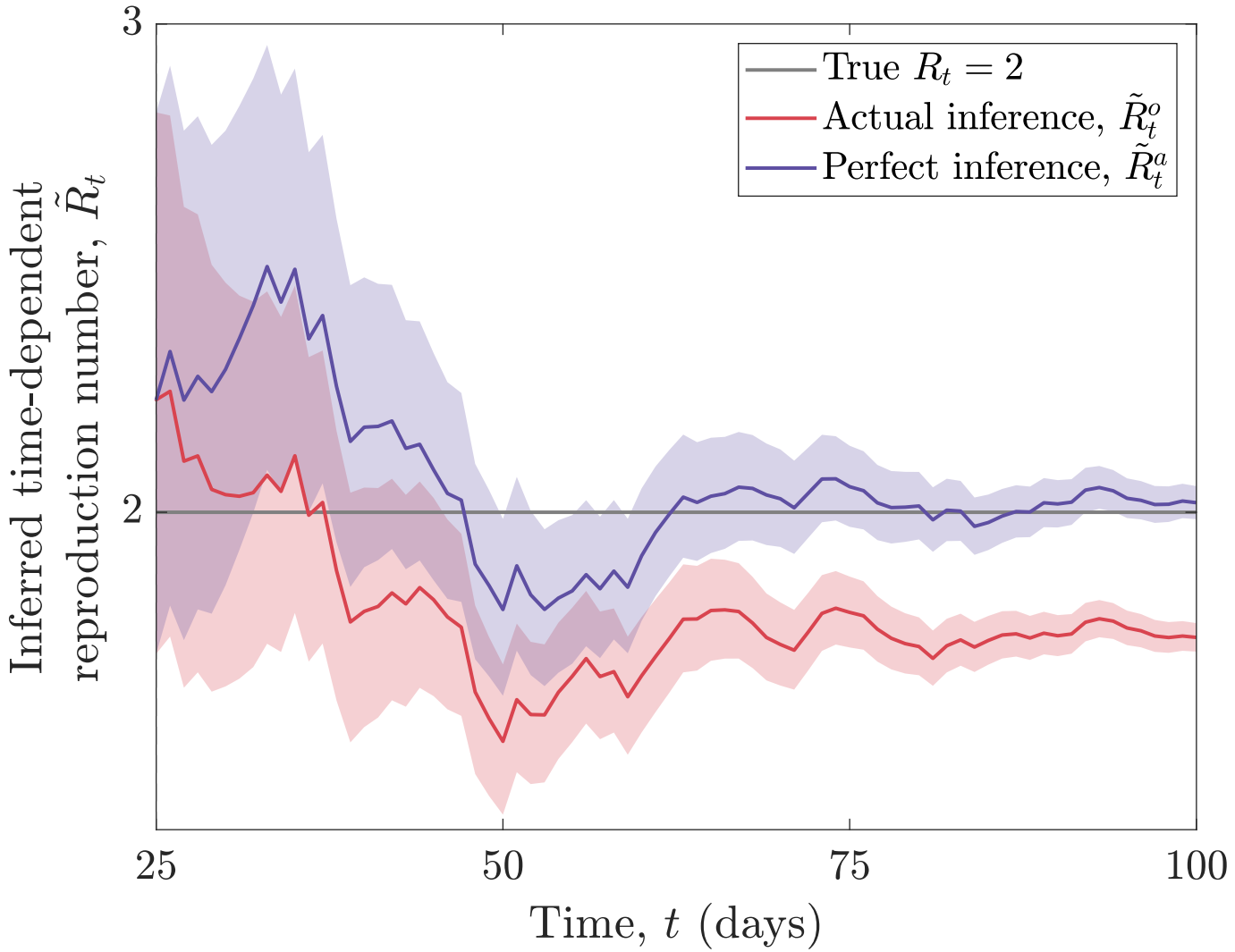
Best-case and realistic R_t estimates may be inferred using the true and recorded generation intervals respectively.



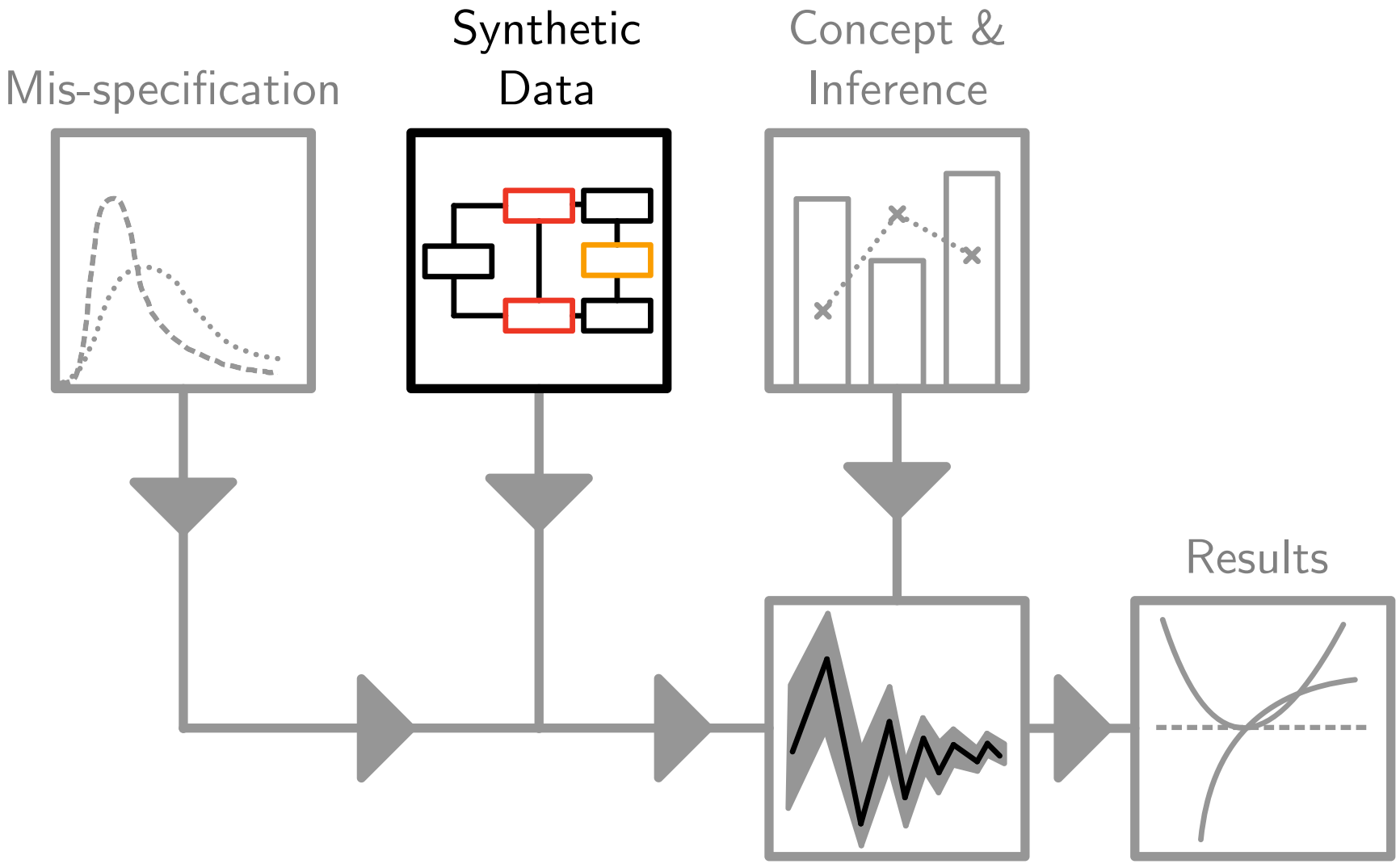
The ratio of these two inferences can be made, $\tilde{R}_t^o / \tilde{R}_t^a$



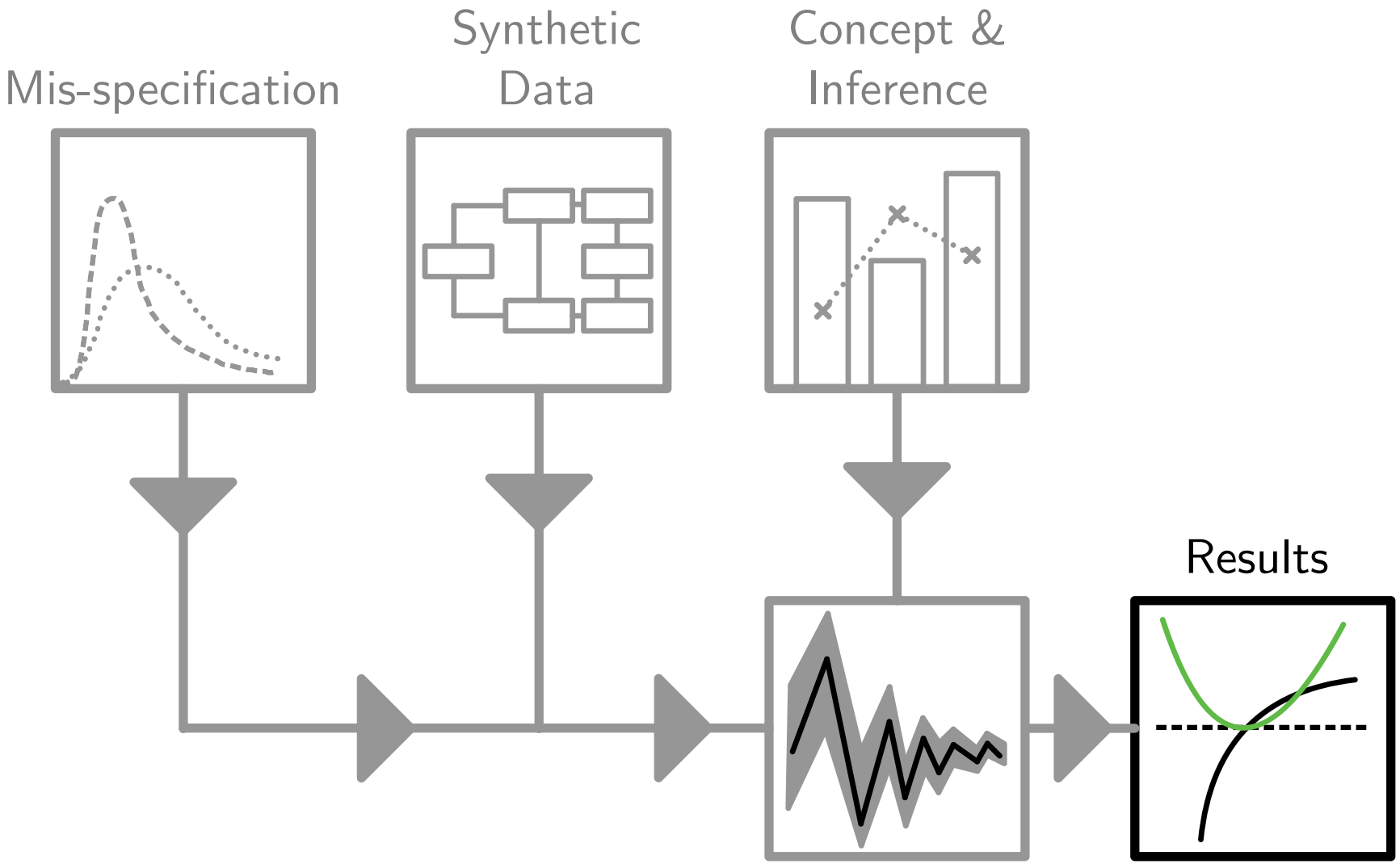
Systematic over/under estimation bias does not imply over/under R_t estimation



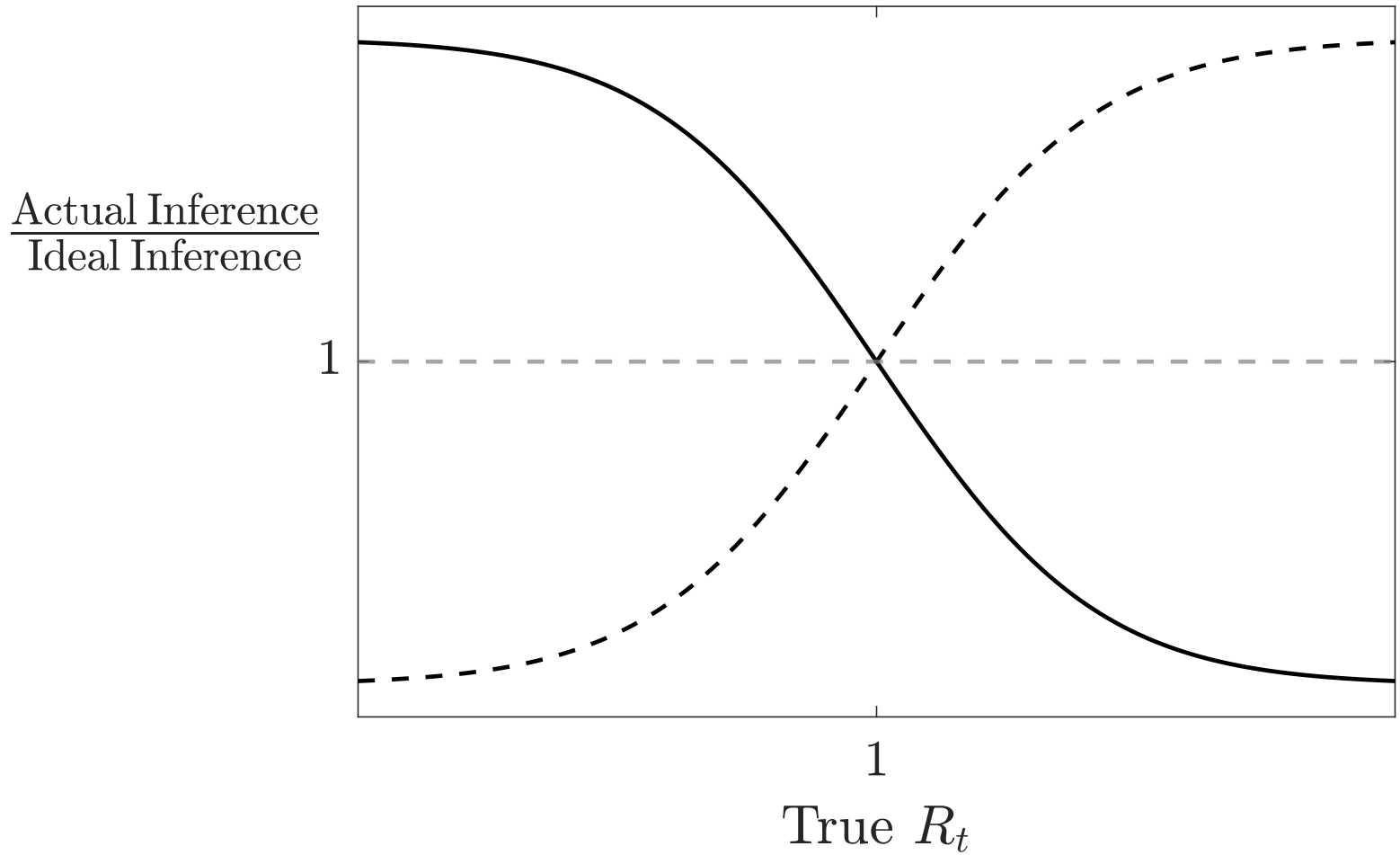
Comparing perfect & realistic inferences via synthetic data



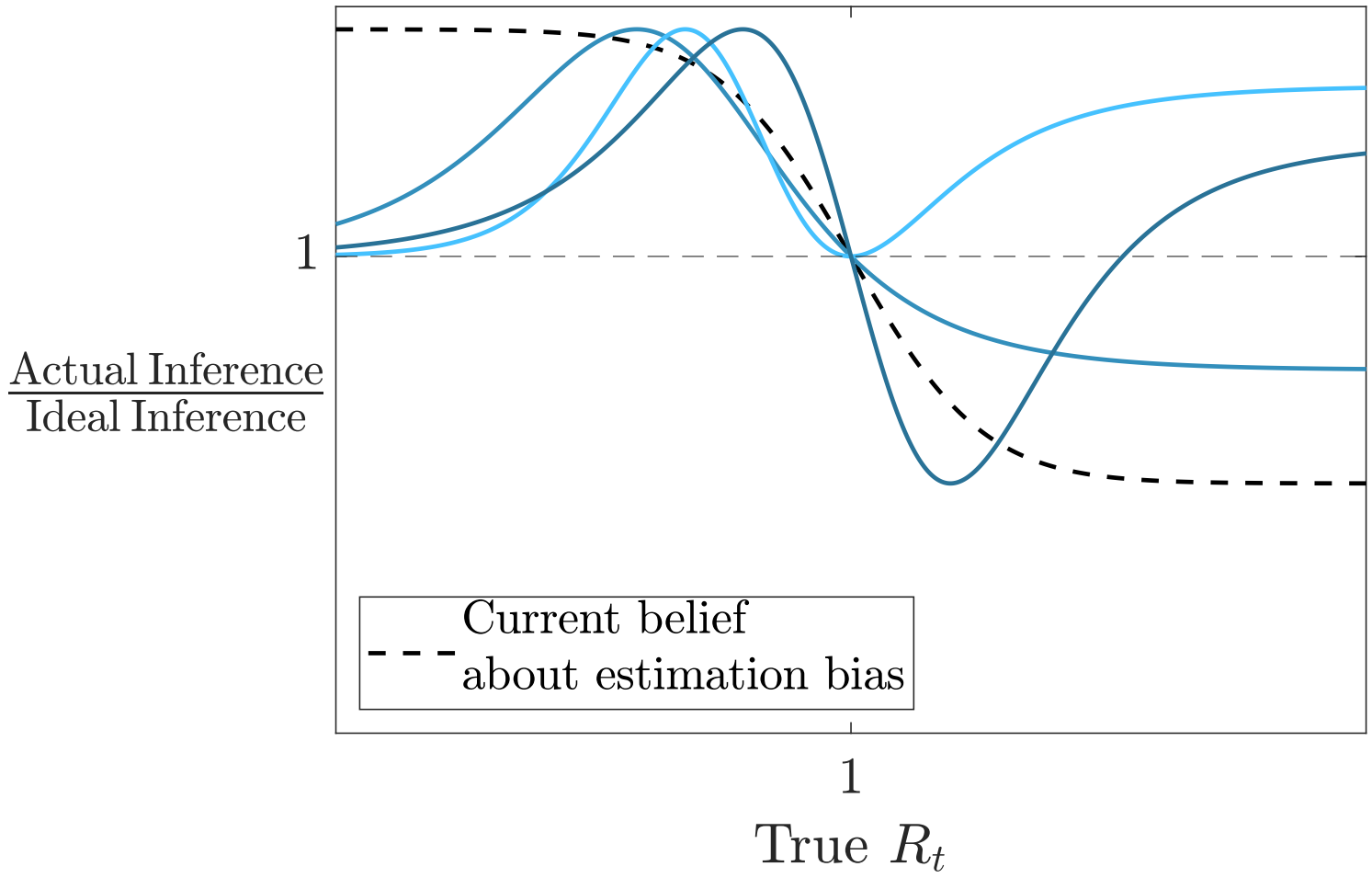
Results


Mis-specification
Synthetic Data
Concept & Inference
Results

The consensus view is that R_t inference is more bias, the greater R_t is from 1



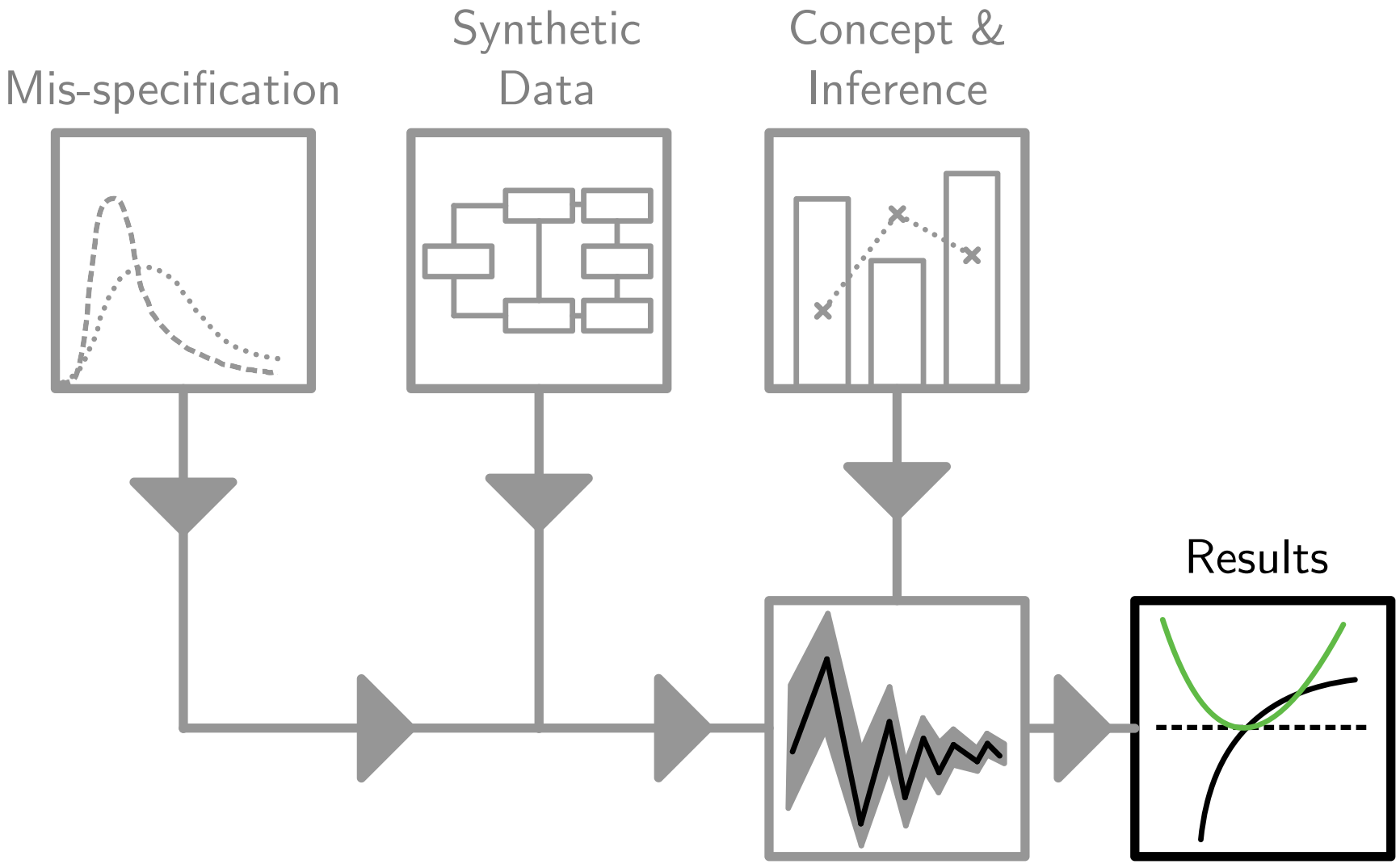
Our results show that this view is false: Estimation bias is more complex



Using a limited knowledge of issues in data collection, we can comment on what type of bias may be present

Source: https://www.gov.uk/guidance/the-r-value-and-growth-rate, 25/09/21

Results


Mis-specification
Synthetic Data
Concept & Inference
Results

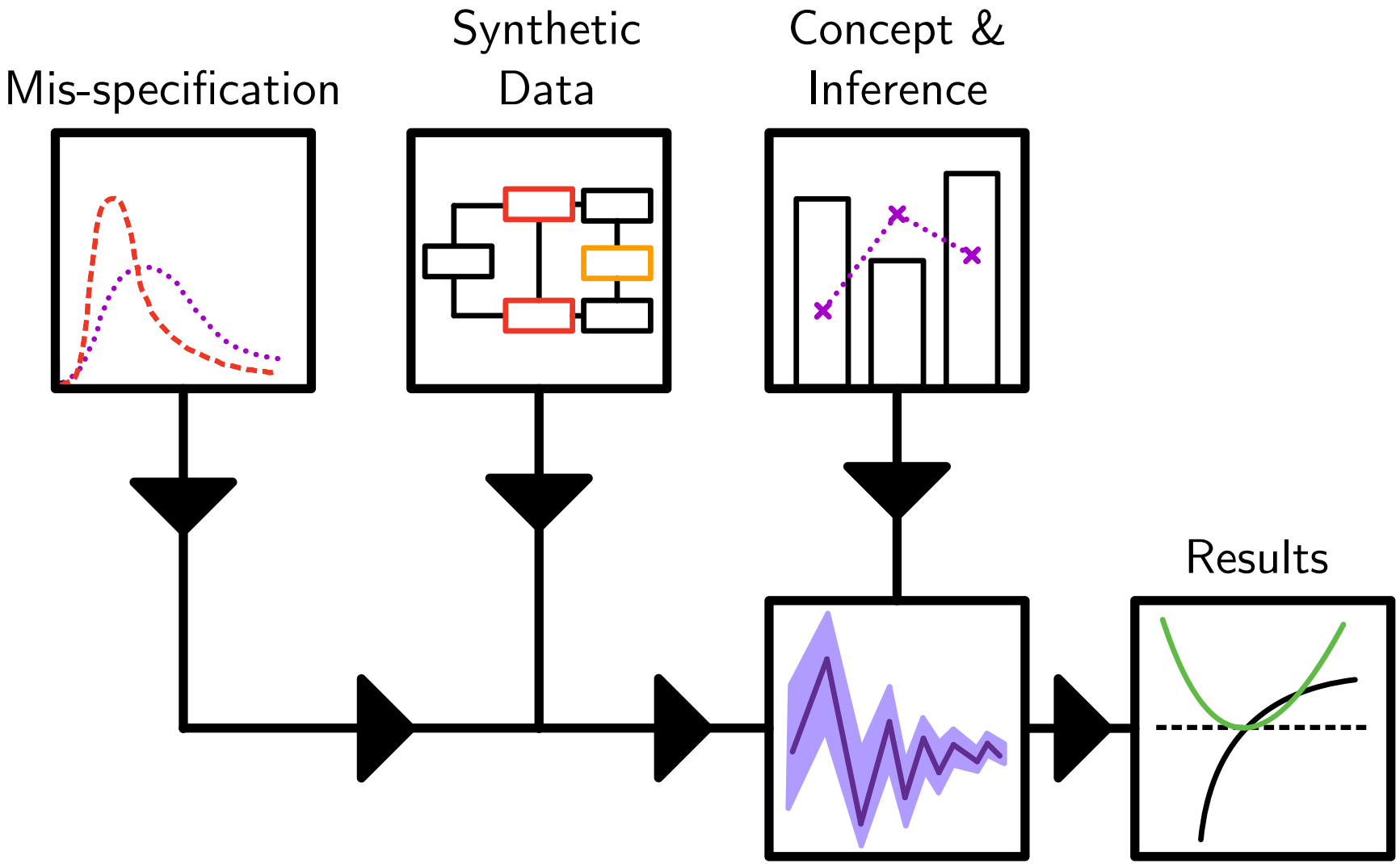
The team!

Zak Ogi-Gittins

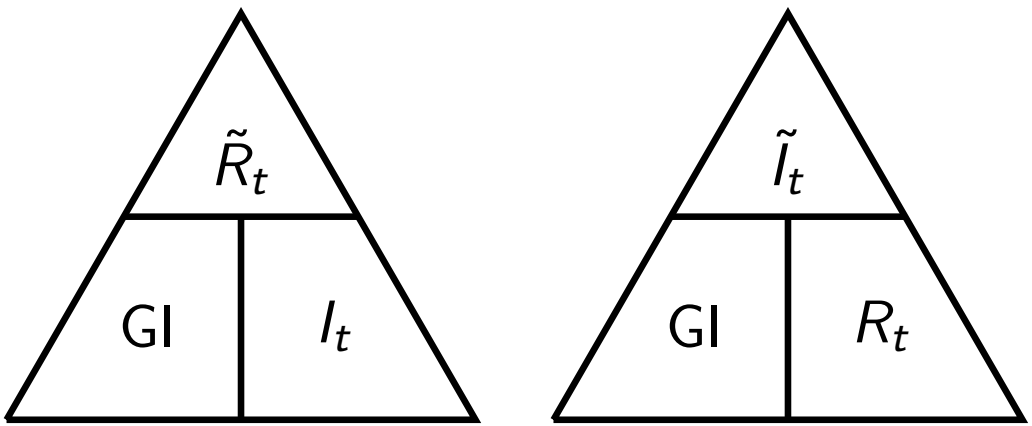
Robin Thompson

Ed Hill

Thank you for listening, any questions are welcome!



Incidence time-series can be synthetically generated using the same method reversed



The $\tilde{X}$ notation indicates that X is either estimated or synthetically generated. main