Highlights

We compared methods of collecting information from multiple infectious disease models.

Key epidemic characteristics were less well represented by a quantile-summary method.

Collecting trajectories enabled continuous evaluation against newly observed data.

The significance of information gain or loss varies with each collaboration’s aims.

Collaborative comparisons among infectious disease models of the future are increasingly common. In the process of collecting multiple projections of the future, such collaborations may gain or lose relevant information. We compared two methods of collecting information from multiple models provided to the European COVID-19 Scenario Hub. We compared a method that summarised model information at each time-step, against collecting raw model output of simulated trajectories over time.

The summary method was less representative of key epidemic characteristics, such as the magnitude and timing of epidemic peaks. We also explored whether a single collection of model outputs can provide continuously useful information about an evolving epidemic over time, with potential for further work to investigate when this would be most useful.

The value of different information gains and losses likely varies with the aims of each collaborative effort. Comparative collaborations should consider the requirements and flexibility of projection users when deciding on a method of model collection.

*Background.* Collaborative comparisons and combinations of epidemic models are used as policy-relevant evidence during epidemic outbreaks. In the process of collecting multiple model projections, such collaborations may gain or lose relevant information. Typically, modellers contribute a probabilistic summary at each time-step. We compared this to directly collecting simulated trajectories. We aimed to explore information on key epidemic quantities; ensemble uncertainty; and performance against data, investigating potential to continuously gain information from a single cross-sectional collection of model results.

*Methods.* We compared July 2022 projections from the European COVID-19 Scenario Modelling Hub. Five modelling teams projected incidence in Belgium, the Netherlands, and Spain. We compared projections by incidence, peaks, and cumulative totals. We created a probabilistic ensemble drawn from all trajectories, and compared to ensembles from a median across each model’s quantiles, or a linear opinion pool. We measured the predictive accuracy of individual trajectories against observations, using this in a weighted ensemble. We repeated this sequentially against increasing weeks of observed data. We evaluated these ensembles to reflect performance with varying observed data.

*Results.* By collecting modelled trajectories, we showed policy-relevant epidemic characteristics. Trajectories contained a right-skewed distribution well represented by an ensemble of trajectories or a linear opinion pool, but not models’ quantile intervals. Ensembles weighted by performance typically retained the range of plausible incidence over time, and in some cases narrowed this by excluding some epidemic shapes.

*Conclusions.* We observed several information gains from collecting modelled trajectories rather than quantile distributions, including potential for continuously updated information from a single model collection. The value of information gains and losses may vary with each collaborative effort’s aims, depending on the needs of projection users. Understanding the differing information potential of methods to collect model projections can support the accuracy, sustainability, and communication of collaborative infectious disease modelling efforts.*Data availability.* All code and data available on Github: <https://github.com/covid19-forecast-hub-europe/aggregation-info-loss>