Optimal algorithms for controlling infectious diseases in real time using noisy infection data

Sandor Beregi[1] and Kris V. Parag[1]

[1]*Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, United Kingdom*
For correspondence: (k.parag@imperial.ac.uk)

Abstract

Deciding when to enforce or relax non-pharmaceutical interventions (NPIs) based on real-time outbreak surveillance data is a central challenge in infectious disease epidemiology. Reporting delays and infection under-ascertainment, which characterise practical surveillance data, can misinform decision-making, prompting mistimed NPIs that fail to control spread or permitting deleterious epidemic peaks that overload healthcare capacities. To mitigate these risks, recent studies propose more data-insensitive strategies that trigger NPIs at predetermined times or infection thresholds. However, these strategies often increase NPI durations, amplifying their substantial costs to livelihood and life-quality. We develop a novel model-predictive control algorithm that optimises NPI decisions by jointly minimising their cumulative, future risks and costs over stochastic epidemic projections. Our algorithm is among the earliest to realistically incorporate uncertainties underlying both the generation and surveillance of infections. We find, except under extremely delayed reporting, that our projective approach outperforms data-insensitive strategies and show that earlier decisions strikingly improve real-time control with reduced NPI costs. Moreover, we expose how surveillance quality, disease growth and NPI frequency intrinsically limit our ability to flatten epidemic peaks or dampen endemic oscillations and why this potentially makes Ebola virus more controllable than SARS-CoV-2. Our algorithm provides a general framework for guiding optimal NPI decisions ahead-of-time and identifying the key factors limiting practical epidemic control.

Introduction

When and how should we intervene in order to most effectively manage an emerging infectious disease? This is a question that is at the core of public health policy-making and has been the subject of ongoing debate [1, 2]. This decision problem is especially crucial during the early stages of an outbreak when there is no or limited immunity in the population and vaccines or other pharmaceutical remedies are unavailable. In this situation, the main control measures are non-pharmaceutical interventions (NPIs), such as mandatory social distancing, mask-wearing, lockdowns and travel restrictions [3, 4, 5].

Outbreak management policies need to balance the risks from mistimed and ineffective intervention decisions with the likely costs of those decisions. An NPI that is applied too slowly or removed too quickly risks large epidemic peaks or rebounds that overburden healthcare systems [6]. However, more conservative approaches may prompt long periods of restrictions that incur costs due to closed economic sectors and borders as well as limited mobility [7].

Optimising the counteracting costs and risks of NPIs is a challenging and enduring problem. This problem is further exacerbated by the practical constraints of real-time surveillance. The data available for an unfolding epidemic are subject to multiple sources of noise and uncertainty that fundamentally limit our ability to infer the state of the epidemic [1, 8]. Solutions therefore require evidence-based research into the benefits, risks and societal costs of different NPIs and public health policies [9, 10] as well as rigorous algorithms that can integrate outcomes of that research with uncertain knowledge of the epidemic state to guide decision-making.

Here, we focus on the latter issue and investigate how optimal, data-driven policies can be derived from real-time surveillance data. We leverage ideas from control theory and reinforcement learning and expose exactly how uncertainties in practical surveillance intrinsically limit the optimal policies. This approach, which uses feedback control, dynamically updates NPI choices by feeding back data on the incidence of new infections that should reflect the most recent epidemic state. However, the reporting of the incidence of infections is subject to delay and under-ascertainment

that is often inherent to real surveillance systems. Under-ascertainment of infections can result from asymptomatic and mild infections, which are rarely observed, or from testing capacities [11]. Consequently, we only receive reports on a random fraction of all new infections. Delays can emerge from the lag between infection and symptom onset or confirmation as well as latencies in testing and processing test results. The consequence of this is that the reported time series of cases (or a related proxy for infections) are stochastically behind the actual incidence [1, 12, 13].

These sources of noise and uncertainty sparked an ongoing debate on what is the best approach for controlling epidemics in real time. At least three challenges have influenced this debate. First, although feedback control is widely used to solve real-time problems in electrical and mechanical engineering [14, 15, 16], these strategies can become destabilised by noise and delays [1, 17, 18, 19, 20]. Second, the timing of public health interventions is critical to their efficacy and hence their associated risks and costs [21]. Last, integrating costs, risks and noise within a framework is difficult and often intractable for deriving insights.

In view of these challenges, some studies have proposed feedback-independent methods that are insensitive to noise and uncertainty in real-time epidemic surveillance. One such approach is to implement a pre-set sequence of cyclic switching between lockdowns and periods with no restrictions. This was proposed as a strategy to exit full-lockdown more reliably [22].

Other works have focussed on optimising the timing of specific interventions considering a 'one-shot' control with the start and ending time to be optimised [23, 24], highlighting the importance of timing for efficiency and the detrimental effect of delays in intervention. This makes a case for using evidence-based policies that dynamically optimise interventions to available-data. Some recent studies support this optimal timing approach informed by real-time data to feedback-independent methods, even for uncertainty in data. However, those works do not consider the intrinsic stochasticity of the epidemic and the related costs of interventions [25].

The last challenge, which stems from the complexity of the decision-making process given the uncertainties in data, transmission details and the likely effect of actions, has meant that the majority of approaches in the field on cost-optimal control only consider deterministic models or limited modelling of noise. As a result, there is scope for fully stochastic but rigorous decision-making and modelling frameworks that can guide interventions by providing insight into how cost-optimal choices and various uncertainties interact.

We consider the real-time control problem in a probabilistic setting, where the epidemic is modelled by a renewal branching process [26]. This is more realistic than the deterministic approach, generalisable to multiple diseases and reflects on the intrinsic variability of infections between individuals. We parameterise our renewal models to describe the dynamics of epidemics of COVID-19 and Ebola virus disease. We propose a model predictive optimal control strategy that balances the costs of NPIs against those generated from the infections projected to occur under the renewal process given our NPI choices. Our control approach is based on real-time incidence data which is delayed and under-ascertained, and incorporates the stochasticity of the epidemic generation process. We also incorporate other limiting factors of real-time control, such as constraints on how frequently NPI policies can be changed.

We assess what limitations data quality imposes on the viability of real-time feedback control for epidemic management and how this is influenced by disease growth dynamics. We compare the performance of our proposed optimal control algorithm with two benchmark control strategies that apply decisions based on chosen thresholds or times. We demonstrate that our algorithm not only outperforms these approaches but can adapt to unexpected changes such as the emergence of new variants or reduced effectiveness of NPIs due to behavioural changes.

Results

Optimal epidemic control

Our study focuses on epidemic control at the early stages of an epidemic with no or limited immunity in the population and without any available pharmaceutical remedies. Consequently, the main control measures are NPIs, such as social distancing, mask-wearing, stay-at-home orders, business closures and travel restrictions. These measures limit disease transmission with the aim of reducing the likely numbers of severe infections below healthcare capacities and minimising expected morbidity and mortality. However, these benefits must be balanced against the costs induced by those NPIs, which may include economic downturns and loss of livelihood.

An ideal control policy would keep the incidence of new infections at a manageable (target) level, optimally balancing the costs of treating infections and implementing interventions. However,

this is non-trivial both because disease dynamics can change in real time and our ability to track those changes is strongly limited by the quality of available surveillance data. To achieve this, we propose a *model predictive control (MPC)* framework for optimising epidemic interventions based on real-time incidence data. MPC utilises a mathematical model to project the dynamical behaviour of the controlled system [27].

We outline our MPC approach in panel (a) of Fig. 1. Broadly, this algorithm compares historical incidence data to desired target levels and decides appropriate control actions that minimise both expected future costs and drive us closer to our target. This target is essentially the level of infections policy-makers may believe sustainable. For example, this may be set so that the proportion of severe infections leading to hospitalisation never exceeds healthcare capacities or to ensure a manageable endemic level of infections. This joint target-cost optimisation process is iteratively done in real-time via the feedback loop in Fig. 1 and makes use of short-term projections.

Our algorithm is analogous to a Markov Decision Process [28] and involves an agent (policy-maker) that decides which NPI to implement from a finite action-space of possible NPIs, based on the projected discounted reward. This reward is compounded over a fixed time horizon for each possible action or action-sequence, as explained in the highlighted incidence curve in panel (a) of Fig. 1. In our simulations, we consider three possible actions which we will refer to as full lockdown, limited social distancing and no restrictions. These interventions are modelled as multiplicative reductions in transmission. The reward function is derived from the costs associated with infections and interventions and is evaluated using projections informed by daily new cases data (C_t) and the expected changes in transmission due to our possible intervention choices.

We use cases as we rarely know the actual number of infections I_t due to surveillance imperfections such as under-reporting and delays. The overall cost consists of the cost attributed to the interventions, a penalty term that is proportional to the error between the incidence and our target and an additional penalty term applied for overshoots larger than 50% of the incidence target. The intervention cost is accumulated daily, with large, moderate and no cost for each day under full lockdown, limited social distancing and no restrictions, respectively. At each decision point, the action with the largest projected discounted reward (i.e., the smallest cost) is chosen. We only consider scenarios in which these decision points occur at weekly or larger spacings to model practical review periods.

We focus on three key metrics to evaluate the performance of control algorithms informed by epidemic data that is subject to different sources of surveillance noise. We look at the peak (maximum) of the incidence curve, the bounding envelope (difference between the maximum and minimum) of incidence values in periods when the epidemic is stable (akin to the endemic state) and the cost spent on interventions. According to these metrics a high-performing control algorithm prevents large epidemic peaks, and forces infections into a manageable steady state with small fluctuations without overly applying NPIs. The last is achieved by optimising timing, switching and duration within our available NPI set. The mathematical formulation of our MPC approach is presented in the Methods.

Alternative control strategies

We compare the performance of the above proposed optimal control algorithm with two simpler control strategies: an *event-triggered* feedback control and a cyclic *time-triggered* control strategy. These strategies represent two fundamental approaches for controlling epidemics in real time, which have either been applied or proposed earlier. Note that for all strategies that we consider, we allow tuning so that the strategies can stabilise the observed incidence.

The event-triggered control applies or relaxes lockdowns whenever reported incidence crosses a predefined threshold, which is often heuristically set in practice. This crossing constitutes an event. We illustrate this approach in panel (b) of Fig. 1 for an example incidence time series. Event-triggered control approaches have previously been used to enact interventions, e.g. for influenza [29] and were considered for triggering NPIs to suppress COVID-19 in the UK [3]. Although this strategy applies limited feedback based on the most recent incidence, it is unable to leverage the information in the full past time series or assess the likely future outcomes of its decisions.

In contrast, the cyclic time-triggered control strategy (see panel (c) of Fig. 1) implements a predefined sequence of actions, which is not based on any direct feedback from the epidemic dynamics. The periods with full lockdown or no restrictions can be arbitrarily long, e.g. a 20/10 cyclic control strategy repeats 20 days of full lockdown followed by 10 days of no restrictions. This strategy was proposed as an effective means of COVID-19 control when surveillance data are poor quality and hence unreliable for informing decisions [22].

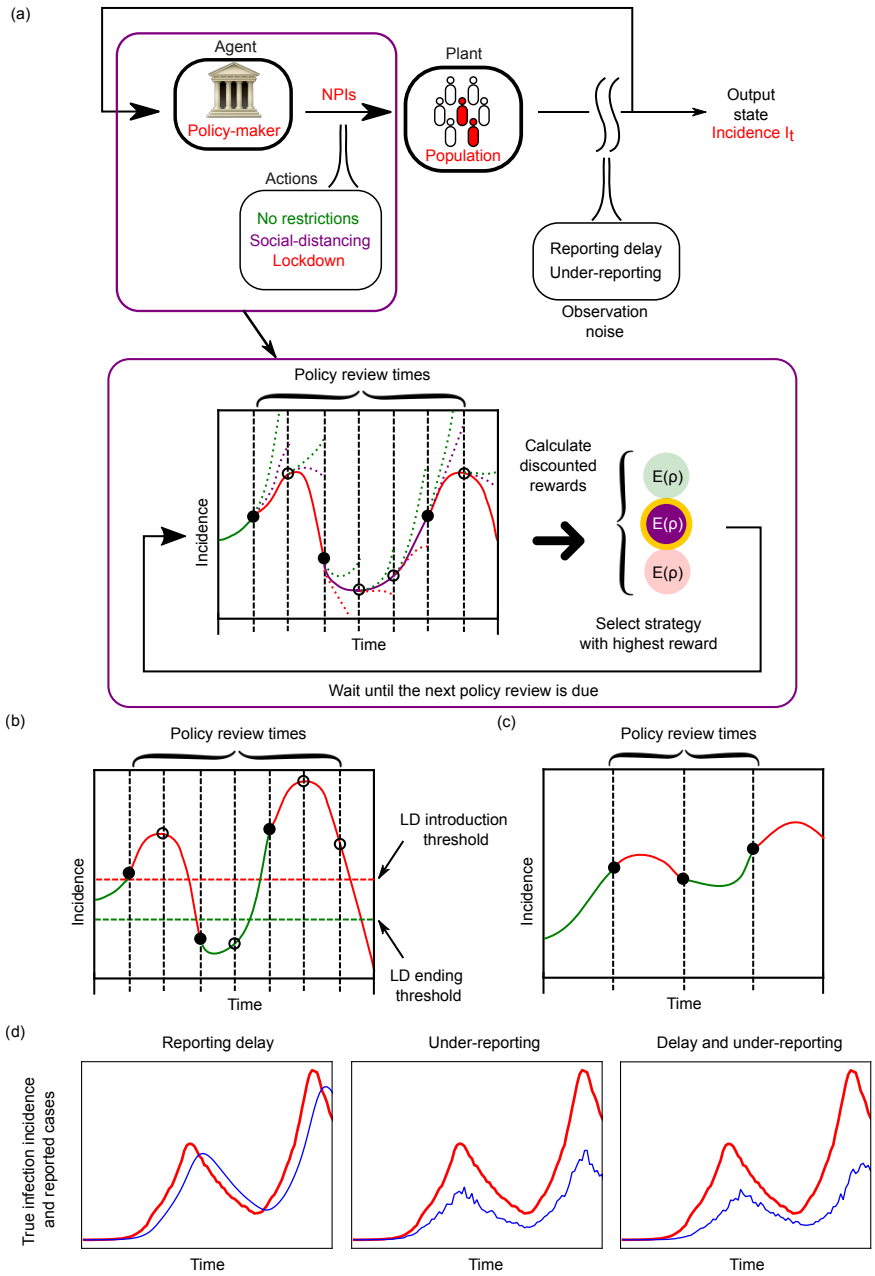

Figure 1: Panel (a): Schematic diagram of model predictive control for optimising epidemic interventions. The top chart introduces the elements of the feedback loop where the actions of the agent are chosen according to the incidence of new infections (or a proxy such as new cases) which is the monitored output state. The highlighted panel explains the model predictive method for selecting the optimal NPI from the action space which is based on using short-horizon projections and the expected reward, $E(\rho)$, under those projections for various strategies. The strategy with the highest reward (minimum cost) is implemented. The reward or cost here usually depends on how far the epidemic state is from our desired objectives, which may aim at jointly reducing both severe epidemic outcomes and intervention intensity. Panels (b) and (c) show event-triggered and time-triggered alternative strategies, with NPIs implemented or relaxed based on incidence thresholds (the event trigger) or according to a pre-defined schedule (the time trigger). Panel (d) illustrates how realistic surveillance imperfections such as reporting delay and under-reporting distort the true incidence of infections into the incidence of cases, which we practically must use to inform decision-making. Blue curves represent the reported cases, while red curves indicate true infection incidence. Reporting delay manifests (approximately) as a time-lag with respect to the true incidence curve, whereas under-reporting results in a stochastic downscale of the incidence curve along the vertical axis.

Observation noise and uncertainty

Infection data are rarely observed directly. As a result, a proxy such as the incidence of reported cases is commonly used for informing decisions in real time. Cases can be modelled as noisy versions of infections [30, 31] that are subject to under-reporting and delays in reporting. We examine how the performance of the above control strategies varies with the levels of these noise sources.

Reporting delay describes the time between an infection and when it is reported as a case. This subsumes lags between infection and symptom onset, onset and presentation at a health facility and processing time to confirmation (e.g., via testing). As explained in panel (d) of Fig. 1, the reporting delay is perceived as a probabilistic shift in time of the reported cases when compared to the actual infection data.

Under-reporting or under-ascertainment occurs when not all infected individuals are tested and reported as cases. This may result, for example, when infections are asymptomatic or only cause mild symptoms so that infected individuals do not seek medical care or testing. Under-reporting can be modelled as a (stochastic) downscale of the infections i.e., a random fraction of new infections appear as new cases.

In the following sections we explore optimal MPC, event and time triggered control strategies and their performance as under-reporting and delays increase. Full details (including mathematical descriptions) of the algorithms we employ are available in the Methods.

Optimal MPC performance

We demonstrate the performance of our MPC algorithm first on perfectly observed incidence data and then explore the influence of practical surveillance limitations. Figure 2 presents four scenarios using MPC to mitigate an infectious disease with generation time (mean: 6.5 days [32, 33]) and basic reproduction number ($R_0 = 3.5$ [34]) chosen to match those previously estimated for COVID-19. Row (a) shows the ideal case where the agent has access to the true incidence of infections. Here, the MPC strategy is able to keep incidence near the target incidence level (5000 new infections/day). Note that we cannot precisely track the target even in this ideal data setting due to intrinsic stochasticity in the epidemic, a finite action space of possible interventions and a policy review period that is at least a week. The resulting fluctuations about the target are a measure of the fundamental control performance under these settings.

In rows (b) and (c) we demonstrate how reporting delay and under-reporting separately affects MPC (cf. panel (d) in Fig. 1). As indicated by the panels in the right column, surveillance imperfections of infection incidence also show up in the estimated reproduction number leading to discrepancies between the observation and the true state of the epidemics. Row (b) shows the effect of reporting delay with a mean of 7 days. In this case, the MPC strategy is still able to keep incidence at a manageable level, however, the delay in observing cases results in late intervention that spurs a higher peak in incidence and larger fluctuations once the epidemic is under control. The thick highlighted curve shows the cases that inform the agent or decision-maker while the thin highlighted curve are the true (unknown) infections.

Row (c) shows the effect of under-reporting with a mean reporting rate of 0.3, i.e., only 30% of infections are reported as cases. Our MPC algorithm is able to achieve the target level but the true incidence fluctuates at a higher level due to the under-reporting noise process which, on average scales down infections to reported cases according to the mean reporting rate. The stochasticity of the under-reporting also occasionally misleads the MPC strategy to believe that the epidemic is under control and the incidence is below the target level or the effective reproduction number is smaller than its true value, causing higher epidemic peaks than in the ideal surveillance scenario. Interestingly, the average proportion of NPIs across the simulation time horizon is similar across scenarios (a-c), indicating that the deviations in the performance of the optimal MPC strategy are due to mistimed interventions resulting from the noise in surveillance, causing either overly conservative or relaxed policies.

Row (d) in Fig. 2 shows the effect of increasing the policy review period from 7 to 14 days with no observation noise. In this case, the MPC strategy is still able to keep incidence at a manageable level, however, the fluctuations in daily incidence, and consequently the peak and the bounding envelope of the later stabilised epidemic are larger as the agent has a reduced ability to intervene and update control actions. The overall effect of increasing the policy review period is similar to having a reporting delay as ultimately, both lead to delayed responses.

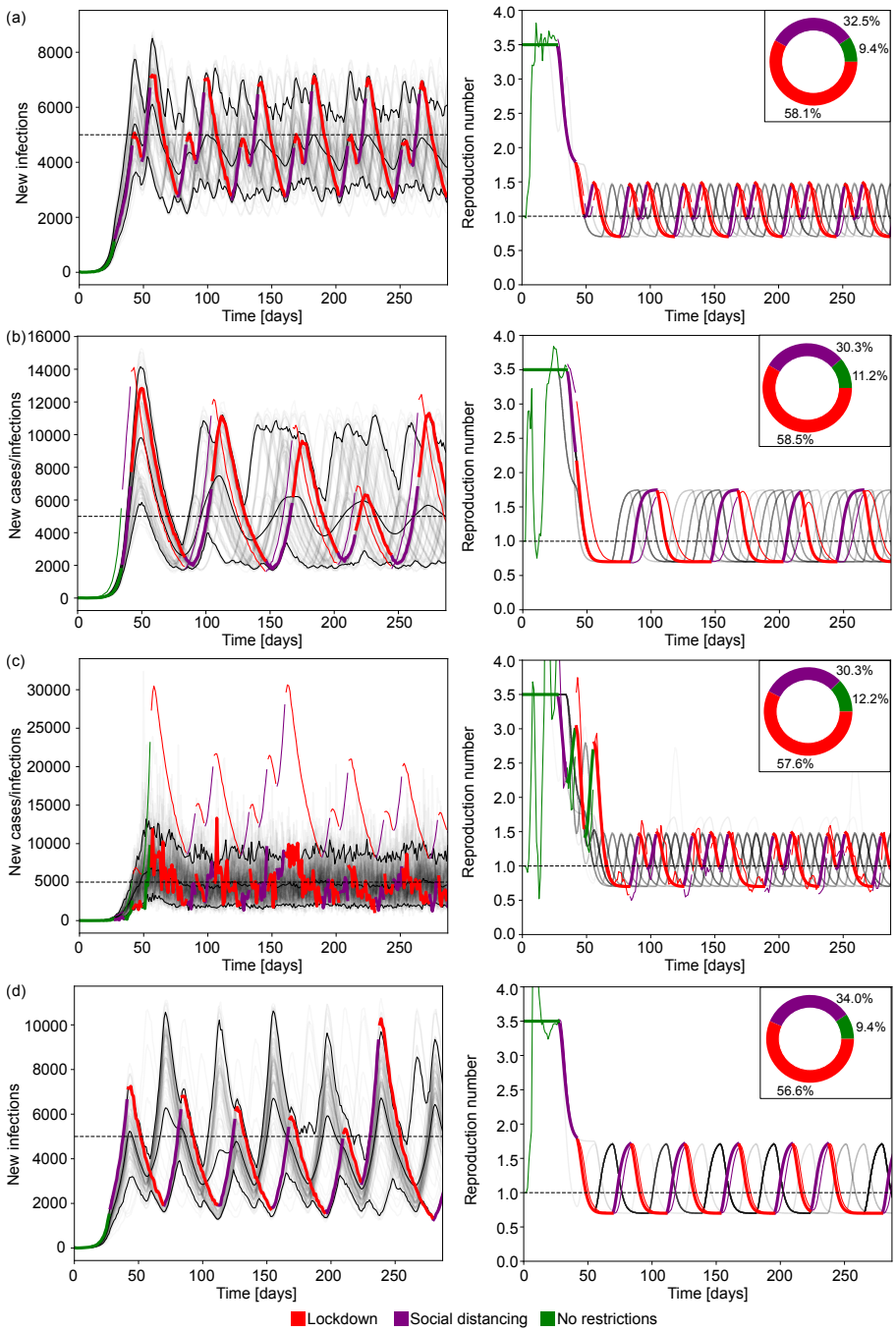

Figure 2: Simulation results with optimal epidemic control. The left panels show reported cases from ensembles of 100 simulations using generation times and reproduction numbers estimated for COVID-19. The faded curves show different individual realisations of the epidemic with the three black curves marking the 5% and 95% percentiles of the ensemble and the mean reported cases. The horizontal dashed line shows the incidence target. The highlighted thick curve of reported cases is coloured based on the NPI implemented on a given day. The coloured thin curve indicates the true incidence corresponding to that highlighted simulation. The right column shows similar diagrams for the effective reproduction number. In that column the faded grey curves and the highlighted curve represents the estimated effective reproduction number from the true infection incidence whereas the thin curve indicates the estimated value from the reported cases in the highlighted realisation of the epidemic. The inset pie charts in the right column indicate the ratio of days spent under a given NPI across the full simulation ensemble. Row (a) represents the baseline case without delay or under-reporting with a policy review period of $t_{\mathrm{rev}} = 7$ days. Compared to the baseline case, panel (b) shows simulations with reporting delay (mean 7 days, shape parameter $\alpha = 5.0$), panel (c) simulations with under-reporting (with mean reporting rate 0.3, dispersion $a = 8.0$). Panel (d) has no observation noise but provides simulations under an increased policy review period of $t_{\mathrm{rev}} = 14$ days.

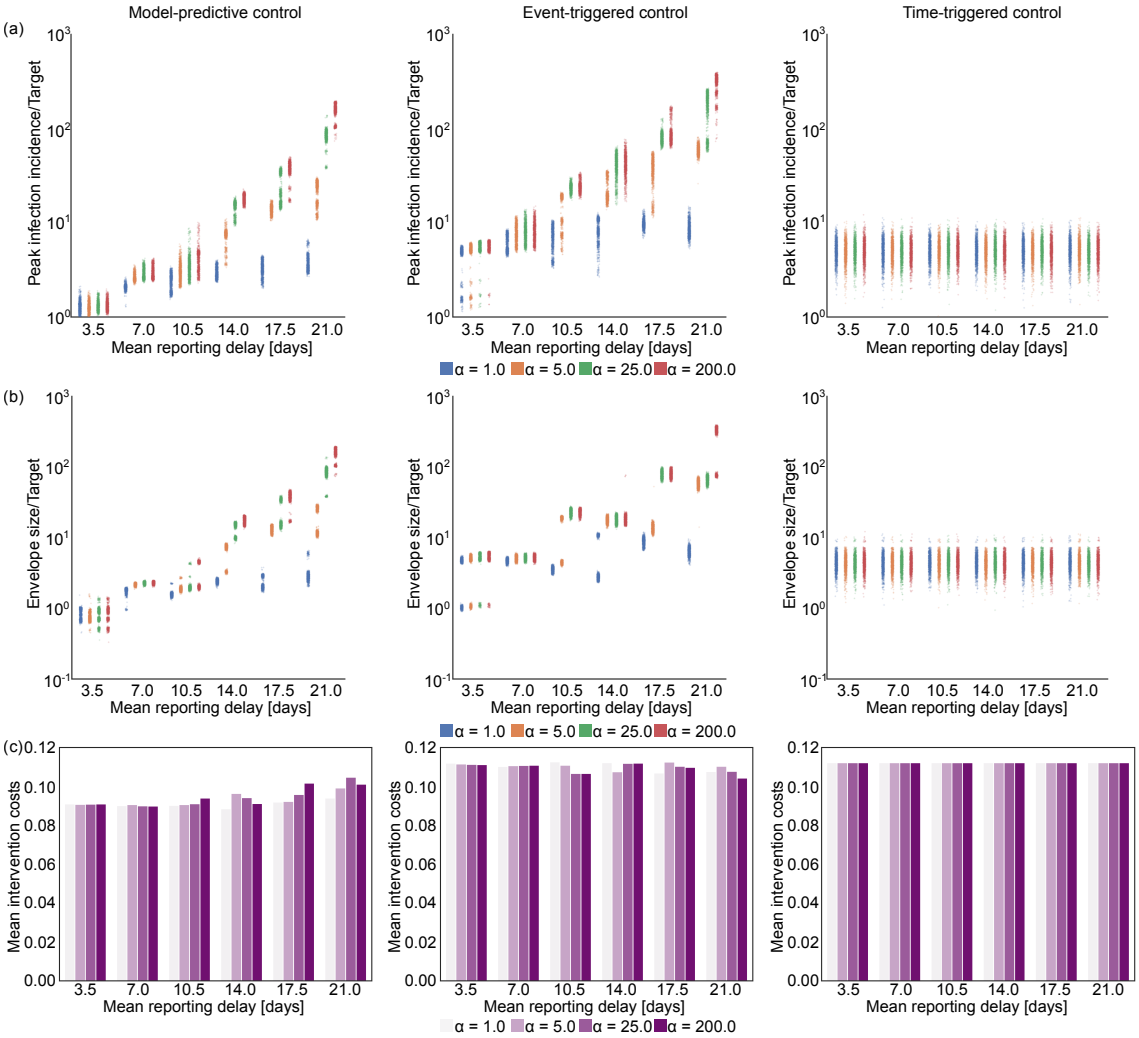

Figure 3: The impact of increasing reporting delay on optimal control. The left panels show results for our MPC algorithm, while the middle and right panels respectively present equivalent outputs under event-triggered (lockdown and relaxation thresholds of $C_{\mathrm{LD}} = 2500$ new cases/day, and $C_{\mathrm{relax}} = 1500$ new cases/day) and time-triggered (cycle of 45 days lockdown, 9 days of no restrictions, starting on day 38) strategies. Row (a) shows scatterplots for peak incidence and row (b) illustrates the steady-state envelope size relative to the target incidence across different time delay distributions. The horizontal axis represents different mean reporting delays with colours depicting different dispersion levels. Larger values of the parameter α indicate more deterministic delays. Row (c) shows the mean intervention costs for each ensemble of 1000 epidemics, simulated under estimated parameters from COVID-19. An equivalent analysis for simulated Ebola virus disease epidemics is presented in Fig. 7 of the Supplement.

The limits of control due to delayed reporting

We assess how the delay in reporting limits the performance of each control strategy. We consider scenarios with different but stationary reporting delay distributions. The delay for a single infection follows a Gamma distribution $\tau \sim \text{Gamma}(\alpha_\tau, \beta_\tau)$, with shape and scale factors α_τ and β_τ, respectively. The mean reporting delay is then $\alpha_\tau \beta_\tau$ while the variance is $\alpha_\tau \beta_\tau^2$. This is a common model of reporting delays and has been used to describe surveillance of COVID-19 and Ebola virus disease among others [32, 35, 36, 37]. To control the mean delay τ_{mean} and dispersion α directly we re-parametrise the distribution by the choice $\alpha_\tau = \alpha$ and $\beta_\tau = \tau_{\text{mean}}/\alpha$. This means that for a given mean reporting delay, the variance is inversely proportional to the dispersion parameter α, i.e., larger values of α correspond to more deterministic reporting delays.

We consider 6 mean reporting delays ranging from 3.5 to 21 days, each with 4 different variances. This allows us to characterise how the mean and the dispersion of the reporting delay limit optimised and heuristic control strategies. For each scenario, we run an ensemble of 1000 simulations, each with a different random seed. Fig. 3 plots key results for simulations under COVID-19 disease parameters. Each column of panels depicts the performance of a different control strategy: MPC, event-triggered, and time-triggered cyclic control. The first row plots the peak incidence, the second row shows the size of the steady-state solution envelope (the range of oscillations after the epidemic is under control) and the third row charts the average costs of NPIs. When calculating the steady-state envelope, we look at the maximum and minimum of daily cases after the incidence of cases falls below the target and the effective reproduction number is below 1. It is not possible to determine a steady-state envelope for every epidemic realisation of every scenario as under some parameter combinations the control algorithm may fail to stabilise the outbreak. If this happens we use the peak value as the size of the solution envelope. We find that the distribution of peaks and steady-state envelopes are multi-modal for certain parameter combinations. This results from applying a controller that has a discrete action space and policy review time longer than the algorithm time step (which is daily).

For the comparison in Fig. 3, we tuned the parameters of the different control algorithms such that they involve similar intervention costs. Additionally, we selected a rolling policy of 45 days of full lockdown followed by 9 days of no intervention (i.e. a 45/9 cyclic policy), which is enacted 39 days after the simulated epidemic started. This is to ensure that the long term average effective reproduction number is approximately 1, so we observe the epidemic reaching a near steady-state. Arbitrary choices of policy lengths may fail to stabilise the epidemic, resulting in continued (but slower) growth or rapid suppression that is costly.

Figure 3 indicates that reporting delays have a detrimental effect to feedback-control strategies i.e., the MPC and event-triggered control approaches, as mistimed action leads to drastically higher incidence peaks and wider steady-state envelopes. In some realisations, the desired target is exceeded by a factor as large as 100. Interestingly and perhaps counter to intuition, the detrimental effect of reporting delay is less if the variance of the reporting delay is high. Low variance means that the reporting delay is almost deterministic (see results with higher values of the dispersion parameter α), while high variance implies that we have a larger portion of cases where the delay is small (and also more cases subject to very large delays, for the same overall mean lag). As a result, high variance delay distributions allow some access to more recent infection information, which can aid decision making. Row (c) of Fig. 3 shows the average costs spent on interventions for each strategy under different delay settings. As expected, we find the optimal control strategy is notably more economical than either the event-triggered or the time-triggered method.

Time-triggered cyclic control is insensitive to the reporting delay as it follows a pre-defined sequence of actions irrespective of the observed incidence trajectory. As a result, while the costs of the interventions are higher than those of MPC, it is possible to devise predefined intervention strategies that outperform optimal feedback-control under large case reporting delays with means of 10.5-14.0 days or more, depending on the delay dispersion. This exposes the limits that surveillance quality can impose on our ability to reactively control an epidemic. Importantly, as long as the dispersion of our reporting delays is sufficiently large ($\alpha \approx 1$) then optimal MPC strategies offer a cost-effective intervention approach for any mean delay.

In the Supplement, we also present the results of the same analysis for Ebola virus (See Fig. 7). We find that the performance deterioration caused by reporting delays follows similar trends to what we observed for COVID-19 in Fig. 3. However, a key difference is that Ebola virus has a longer generation time, which results in slower dynamics, that reduce some of the negative impact of reporting delays. For comparison, with the MPC approach and a near deterministic ($\alpha = 200$) delay distribution with a mean of 21 days, we obtain peak infection incidence 100-1000 times larger than target for COVID-19. In contrast, the same delay only leads to peaks of about 4-5 times the

target for Ebola virus disease.

The limits of control due to under-reporting

Having isolated the influence of reporting delays, we now examine the impact of under-reporting. This is another common and important surveillance imperfection in which not all newly infected individuals are reported as cases leading to under-ascertainment of the true numbers of infections and hence the size of the epidemic. We model the number of daily reported cases with a Beta-binomial distribution $C_t \sim \text{BetaBin}(I_t, \alpha_\nu, \beta_\nu)$, with shape parameters of α_ν and β_ν. The expected number of reported cases is then $I_t\alpha_\nu/(\alpha_\nu+\beta_\nu)$ while the variance is $I_t\alpha_\nu\beta_\nu(\alpha_\nu+\beta_\nu+I_t)/((\alpha_\nu+\beta_\nu)^2(\alpha_\nu+\beta_\nu+1))$. We refer to the ratio of the expected reported cases and the true infection incidence as the mean reporting ratio $\nu := E[C_t/I_t]$ which is constant in time. To directly control the mean reporting ratio ν_{mean} and the dispersion, we choose $\alpha_n u = a$ and $\beta_n u = a(1-\nu_{\text{mean}})/\nu_{\text{mean}}$. This means that larger values of the dispersion parameter a result in a smaller dispersion in reporting rate.

We consider 6 different mean reporting ratios from 0.1 to 0.85 each with 4 different variances. A smaller reporting fraction means larger under-reporting. Fig. 4 plots the performance of the MPC, event-triggered and cyclic strategies in successive columns. For these diagrams, the target incidence is scaled up according to the mean reporting ratio to have comparable results.

We find that as the variance in under-reporting increases the performance of optimal feedback strategies deteriorates. This follows because the under-reported case curve is effectively a stochastic downscaling of the true infection incidence curve. As a result, larger stochasticity in fluctuations for a given mean reporting rate more substantially distorts the observed incidence and misleads control. Low variance means that the under-reporting is more deterministic, so that reported case incidence better resembles the shape of the true infection incidence curve.

Both our MPC and the event-triggered strategy show similar patterns in how the peak and envelope of the controlled epidemics vary with the under-reporting statistics. However, the MPC algorithm achieves better performance with smaller intervention costs. This improvement derives from the the MPC approach leveraging all the available historical information in the incidence curves.

Time-triggered cyclic control is not affected by under-reporting because it is agnostic to the incidence data and how it is reported. The apparent variation in the performance of the cyclic control strategy is due to the scaling of the target according to the under-reporting rate. While under moderate noise and uncertainty, our MPC approach is more effective and cost efficient than either of the heuristic methods, cyclic control can outperform MPC if the reporting rate is very low and the variance is high.

Comparing the effect of under-reporting on control of COVID-19 and Ebola virus (see Fig. 8 in the Supplement) we find similar trends for both pathogens. Even though the drop in performance that we observe is less pronounced than that due to reporting delay, the slower dynamics of Ebola virus still makes it more controllable than COVID-19 when uncertainty in case under-reporting is present. This is evidenced by the solution peaks and steady-state envelopes being distributed in a smaller interval across the simulation ensembles for Ebola virus disease than for COVID-19.

Integrating noise and intervention frequency

Having explored the performance limits induced by reporting delays and under-reporting in isolation, we now consider their combined impact. We analyse ensembles of simulations under realistic noise distributions for COVID-19 and Ebola virus disease. The uncertainty in case reporting data can vary markedly depending on the context and national or regional differences in how surveillance is conducted. In [1], reporting delays of 9-12 days were estimated for COVID-19 in Italy, whereas [38] inferred case-reporting rates between 7-38 % across France. Based, on these, we consider a mean reporting delay of 10.5 days with a dispersion of ($\alpha = 5.0$) and a mean reporting rate of 0.3 with a dispersion of $a = 8.0$.

Rows (a) and (b) in Fig. 5 show results for simulated COVID-19 epidemics under realistic noise distributions and subject to policy review periods of 7 and 14 days, respectively. We find that the MPC strategy stabilises the epidemic around the target, but fluctuations are considerably larger than in the baseline case (see Fig. 2a) due to the delay and under-reporting. This results in an overshoot of about 5-times the target incidence under a 7-day policy review period and around 10-times for a 14-day policy review period. Comparing these results with the simulations in panels (c) and (d) of Fig. 5 for epidemics simulated under Ebola virus parameters, we observe that, the MPC strategy is more effective in controlling these epidemics and the effect of noise is

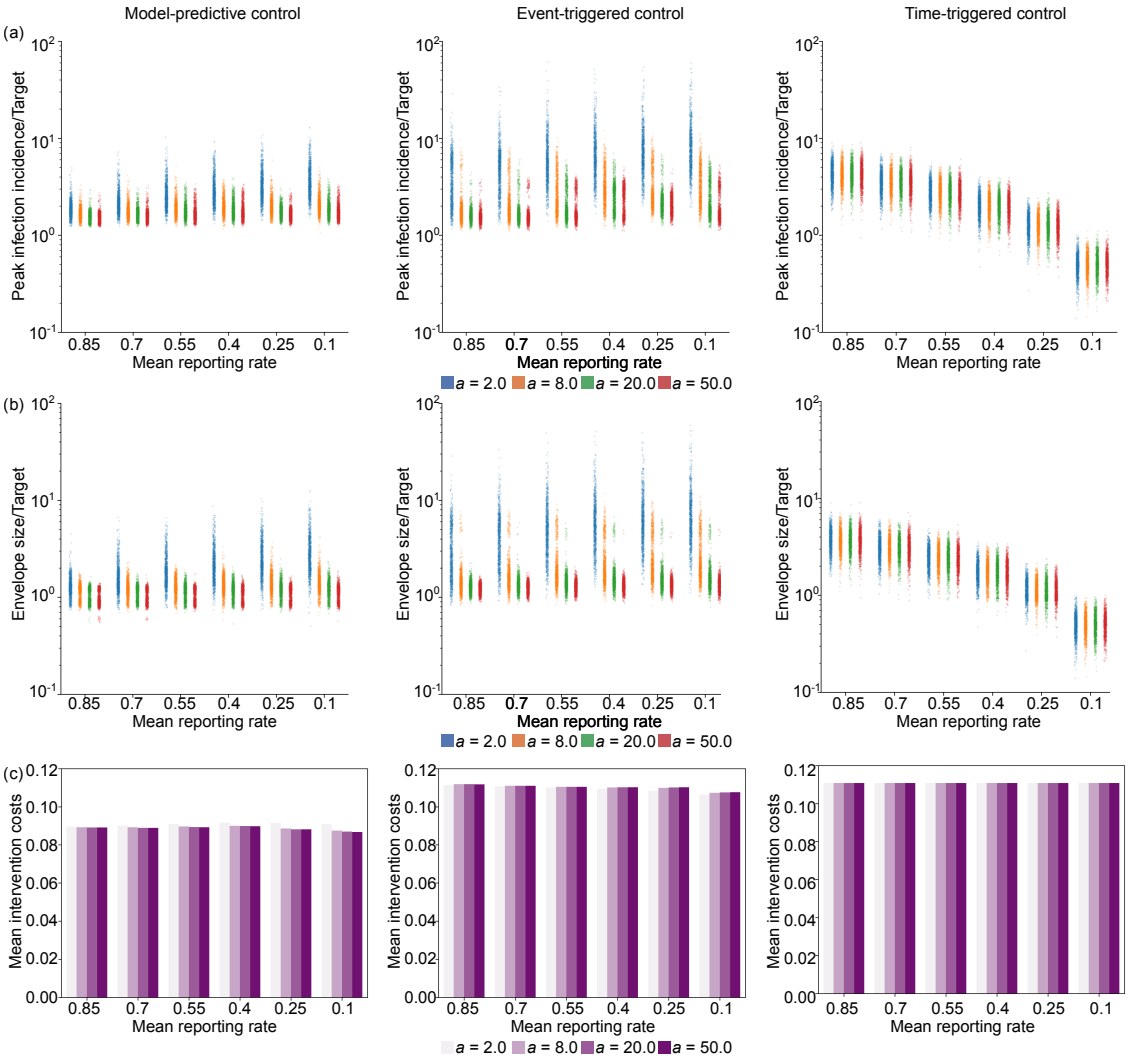

Figure 4: The impact of increasing under-reporting on optimal control. The left panels show results for our MPC algorithm, while the middle and right panels respectively present equivalent outputs under event-triggered (lockdown and relaxation thresholds of $C_{\mathrm{LD}} = 2500$ new cases/day, and $C_{\mathrm{relax}} = 1500$ new cases/day) and time-triggered (cycle of 45 days lockdown, 9 days of no restrictions, starting on day 38) strategies. Row (a) shows scatterplots for peak incidence, row (b) illustrates the steady-state envelope size relative to the target incidence for different reporting rate distributions. The horizontal axis represents different mean reporting ratios with colours depicting different dispersion levels. Larger values of the parameter a belong to more deterministic case reporting distributions (i.e., constant reporting). Row (c) shows the mean intervention costs for each ensemble of 1000 epidemics, simulated under estimated parameters from COVID-19. An equivalent analysis for simulated Ebola virus disease epidemics is presented in Fig. 8 of the Supplement.

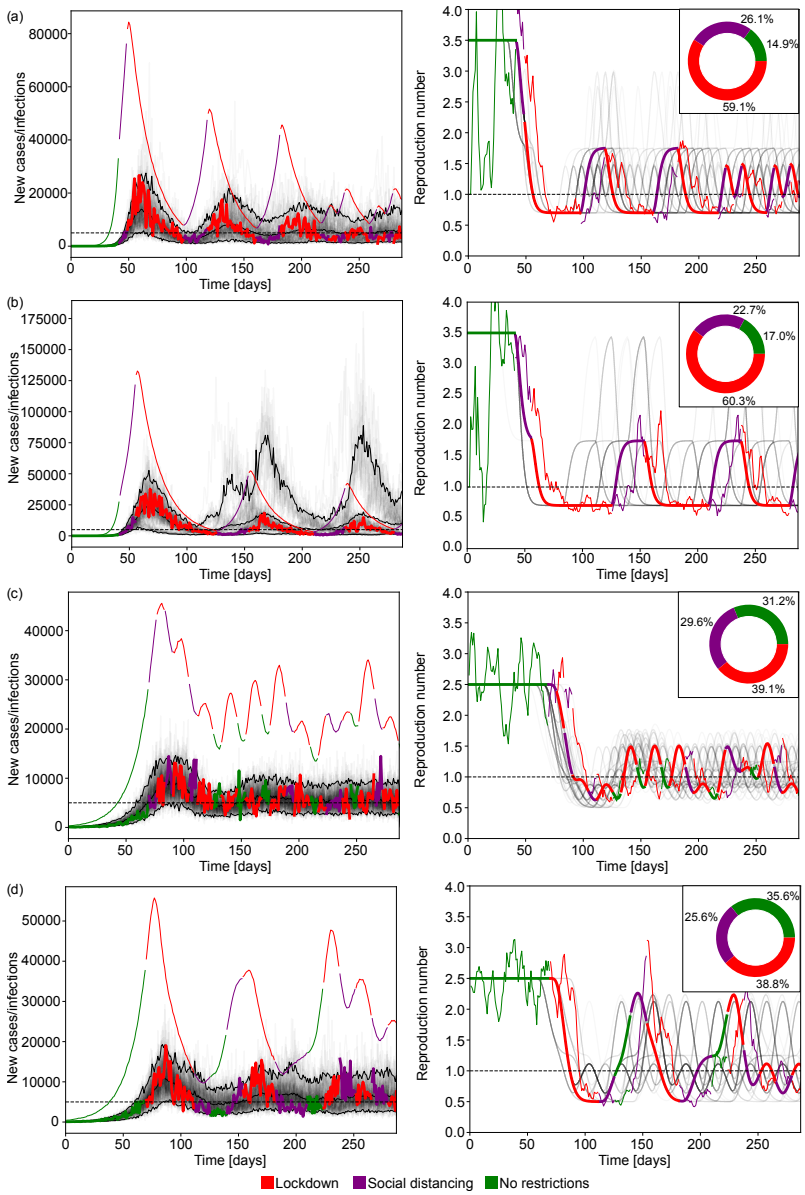

Figure 5: Optimal control and policy review for realistic simulated epidemics. Rows (a) and (b) represent epidemics simulated under a COVID-like generation time and basic reproduction number with 7 and 14 days policy review periods, respectively. Rows (c) and (d) represent epidemics simulated under Ebola-like parameters with 7 and 14 days policy review periods, respectively. Because Ebola virus disease has a longer generation time, we discard a burn in period of 14 weeks to allow the epidemic to grow towards the initial target. The left column show reported cases from ensembles of 100 simulations with mean delay 10.5 days, delay dispersion $\alpha = 5.0$, mean reporting rate 0.25, $a = 8.0$. These settings reflect estimates of realistic surveillance noise from the literature. The faded curves show different individual realisations of the epidemic with the 3 black curves marking the 5% and 95% percentiles of the ensemble and the mean reported cases. The horizontal dashed line shows the incidence target. The highlighted thick curve of reported cases is coloured based on the NPI implemented on a given day. The coloured thin curve indicates the true incidence corresponding to that highlighted simulation. The right column shows similar diagrams for the effective reproduction number. The faded grey curves and the highlighted curve represents the effective reproduction number estimated from true incidence whereas the thin curve indicates the estimated value from the reported cases for the highlighted realisation of the epidemic. The inset pie charts in the right column indicate the ratio of days spent under a given NPI across the full simulation ensemble.

less detrimental to controllability. For the Ebola virus epidemics (panel (c)) we find that running the MPC algorithm with a 7-day policy review period causes a 3-times overshoot of the target incidence of cases, which remarkably, only slightly deteriorates to about 4-times when a 14 policy review period is used (see panel (d)). This occurs because the longer generation time of Ebola virus results in a slower epidemic that allows the MPC algorithm more time to effectively adapt to changes in the epidemic dynamics. Consequently, we must act more swiftly when responding to diseases with shorter generation times as their faster growth can quickly destabilise data-informed policies. Note that in scenarios where the same epidemic parameters were used ((a) and (b) – COVID-19, (c) and (d) – Ebola virus disease) we observe that our MPC algorithm enacted and sustained NPIs in similar time ratios (with COVID requiring more restrictions than Ebola). This confirms that differences in performance are due to timing instead of overly conservative or relaxed strategies.

Discussion

Here we proposed a model predictive control strategy for optimising epidemic interventions that uses incidence data in real time. Our approach is one of the first to design feedback and cost-minimal strategies that integrate both the intrinsic stochasticity of the transmission process and the practical noise that is ubiquitous to real surveillance. Our results indicate that, within the limitations of the data quality, model predictive optimal control is a viable strategy for cost-effectively guiding intervention decisions in real time. Comparing it to earlier reference approaches, the MPC strategy appreciably outperforms both event-triggered feedback control and time-triggered control.

While noise in surveillance data has a detrimental effect on both the MPC and the event-triggered approaches, as both utilise real-time data, within a direct feedback loop, because our MPC approach considers the full epidemic state and projected epidemic dynamics in decision-making, it is able to better leverage the signals within that data. This allows it to simultaneously achieve better or equivalent noise robustness while utilising a smaller intervention budget than the event triggered approach, which only uses feedback relative to fixed thresholds in infection incidence to impose and relax NPIs.

If noise levels are extreme then time-triggered strategies, which schedule NPIs without directly considering real time data, can be more effective in limiting peak incidence in scenarios with large, near-deterministic delays. This marks the limits of data quality for feedback-control strategies. However, the time-triggered strategy also has notable drawbacks because its design requires precise knowledge of both the epidemic parameters and the efficacy of each NPI. Consequently, this strategy is still implicitly linked to infection incidence data in practice. We therefore find strong evidence that the additional complexity, relative to reference strategies, such as event and time-triggered approaches, involved in performing MPC brings substantial advantages. Moreover, because this optimisation is sequential it adapts well to unexpected changes or uncertainties, offering important robustness to the many unknowns during an unfolding epidemic.

For example, if immunity is acquired by infection, then this reduces the susceptible population and decreases the effective reproduction number. In our modelling framework, this can be easily included by setting $R_t = R_{0t}S/N$, where S and N are the susceptible and the total population, respectively. Note that we excluded this from our simulations in order to isolate the impact of the NPIs and to allow fair comparison among epidemics that have dynamics on differing timescales.

The basic reproduction number R_0 could also change due to the emergence of a new strain of the pathogen (e.g., new variants appeared several times during the COVID-19 pandemic [39]). The MPC algorithm has the flexibility to handle these changes in disease dynamics as it infers the reproduction number from data and actively keeps infection incidence around the target level. We present simulations in the Supplement (Fig. 9, panel (a)) in which the basic reproduction number for COVID-19 increased from 3.5 to 4.5 to model the appearance of a new, more transmissible variant. In this case, the MPC algorithm retains control over the outbreak as long as there are actions (i.e., possible interventions) in the action space capable of reducing the effective reproduction number to below 1.

Our MPC algorithm is also capable of accommodating uncertainty in the efficacy of NPIs to reduce the effective reproduction number of the pathogen. As we demonstrate in Fig. 9, panel (b) of the Supplement, whilst uncertainty in the expected effectiveness of NPIs may lead to occasional misjudgements of the optimal action that cause larger peaks in infection incidence, overall the MPC algorithm still effectively controlled the epidemic by actively reacting to the discrepancies between the actual new cases and the target. This result corroborates findings in [25] and highlights why having an adaptive strategy that considers data and action in feedback is beneficial.

Our results also emphasise that while noise can degrade even optimal MPC strategies, these rigorously optimised control approaches are valuable in almost all scenarios. Note that we did not correct for these sources of noise when assessing their detrimental effects on epidemic controllability. While several studies have focussed on estimating and compensating for under-reporting [11] and reporting delays [40, 41, 42], these approaches often require additional knowledge about the reporting process or orthogonal data sources [43]. It is often the case that these are not available or only become available later in epidemics so we preferred to characterise performance under the more practical scenario that little else is known about the epidemic than its time series of cases. Moreover, there are also inherent practical delays involved with the public announcement and implementation of NPIs from the point of decision [44], which would result in delayed action even with perfect knowledge of the epidemic states. For example, having a policy review period that is several weeks for a pathogen with dynamics that vary on daily timescales can mean that interventions are inevitably late or suboptimal. This effectively causes an additional lag in the feedback loop and may itself prevent controllability.

We also found that the speed of disease spread is an important factor that sets the limiting timescales for surveillance and action in order for a set of NPIs to achieve control. It is easier to control a slower spreading pathogen like Ebola virus disease (mean generation time $\sim$ 15 days), as compared to COVID-19 (mean generation time $\sim$ 6 days). Generally, for a slower spreading disease there is more tolerance for having longer reporting times or less frequent policy reviews. Accordingly, there is also less sensitivity to the timing of interventions.

The negative effect of having a lower policy review frequency (or longer time between policy updates) implies that it is ideal to review intervention decisions as often as possible (hence allowing for a more continuous feedback loop). However, as NPIs are intrusive and costly, doing so would probably result in changes in public behaviour that in turn influence the effectiveness of the NPIs [45, 46]. For example, the level of adherence to closure or social distancing policies may wane due to fatigue or perceived risk [47, 48]. Although our MPC approach is adaptive and efficient at controlling unfolding epidemics, it does depend on several assumptions. Specifically, we do not consider how dynamic changes in behaviour, mobility or other individual-level variations in response to policy and epidemic data alter transmissibility. We also assume that the population is well-mixed which means we do not account for heterogeneity in transmission (e.g., superspreading) or spatial and sociodemographic differences that can modulate spread. However, our MPC framework can accommodate some of these sources of heterogeneity (e.g., we can easily include superspreading via more dispersed renewal models) and our future work will focus on better understanding how heterogeneity may affect intervention choices.

While the realism of our study is dependent on having accurate knowledge of costs (both economic and due to health outcomes), our framework is flexible and easily incorporates these as well as finer resolution intervention options (e.g., we can directly expand our action space to include NPIs with intermediate stringency such as mandatory wearing of masks, restrictions to larger meetings, or limiting in-person attendance to work or education). Our goal was simply to construct a general framework and qualitatively assess how optimal and suboptimal but known control strategies depend on realistic surveillance limitations.

Our results unequivocally demonstrate that timing is a crucial factor in intervention efficiency. The same interventions or interventions under the same overall budget applied differently across time can yield markedly different disease control outcomes due to this sensitivity to timing. Even when optimal algorithms such as our MPC approach are applied, mistimed action (due to delays in data or NPI reviews) can be detrimental and substantially reduce policy effectiveness causing high infection peaks. Ascertainment of infections is also important but even suboptimal strategies are more robust to this type of noise than delays. Consequently, improving the speed of epidemic detection and response systems should be a priority for disease surveillance and policy.

Methods

Epidemic governing equations

We model the spread of the disease in a population as a generalisation of the standard renewal branching process [26]. This model is used both to make projections that inform optimal control and to simulate 'ground truth' epidemic trajectories. The renewal branching process is a stochastic model describing how the incidence of new infections on day t, I_t, depends on past infections at

times $s \leq t$ and the characteristics of the disease. This is captured by the Poisson distribution

$$I_t \sim \text{Pois}(\sum_{s=0}^{t} I_s R_s w_{t-s}), \tag{1}$$

where R_t is the effective reproduction number on day t, with the set of weights, w_{t-s} for all s, obtained from the generation time distribution [32] of the disease. We assume that the generation time distribution is known or estimated from other paired transmission data. The weight w_{t-s} is the probability that a secondary infection occurs $t-s$ days after its primary infection. As is standard practice, we model the stochasticity of the generation time with a Gamma distribution

$$t_{\text{gen}} \sim \text{Gamma}(\alpha_{\text{gen}}, \beta_{\text{gen}}). \tag{2}$$

The shape and scale factors α_{gen} and β_{gen} parametrise the probability density function $p_{\text{gen}}(t)$. The weights w_{t-s} used in Eq. (1) are then calculated as $w_{t-s} = p_{\text{gen}}(t-s)/\sum_{k=0}^{t} p_{\text{gen}}(k)$. We consider two generation time distributions, with respective parameters provided in Table 1, which are commonly used to describe epidemics of Ebola virus disease [49] and COVID-19 [33].

The above formulation generalises the standard renewal model, which describes incidence $I_t \sim \text{Pois}(R_t \sum_{t=0}^{t} I_s w_{t-s})$ with the sum $\sum_{t=0}^{t} I_s w_{t-s}$ referred to as total infectiousness of all infectious individuals [50]. However, this classical formula implies that any intervention applied to curb the spread of the disease results in an immediate change in the reproduction number. We apply the generalised formula of Eq. (1) to model scenarios where the infectiousness of a population and the impact of interventions depend on both the history of infections and reproduction numbers. The latter dependence has a smoothing effect that allows for the finite time effects of realistic interventions on transmissibility. A similar generalisation was introduced in [31].

The effective reproduction number is derived from the basic reproduction number R_0, which describes how many people an infected individual is expected to infect in a fully susceptible population. When an NPI is introduced the basic reproduction number R_0 is multiplicatively changed by a factor c_t, yielding

$$R_{0t} = R_0 c_t. \tag{3}$$

We consider an action space with three possible NPIs: no intervention ($c_t = 1$), limited social distancing ($c_t = 0.5$) and full lockdown ($c_t = 0.2$). While this is a simplified classification of NPI types, our control framework allows for more possible actions to be easily modelled. Although in reality the factor c_t is unknown and needs to be estimated, we assume throughout this study, that the effect of any NPI on the reproduction number is known and without uncertainty. Our framework does allow the inclusion of uncertainty on these effects and we present analyses under stochastically varying c_t in the Supplement. However, our goal is not to describe precisely how interventions attenuate transmissibility, but instead to derive insights into how noise influences the optimal timing of interventions generally. Thus, our choices of c_t are meant to be sensible but not exact and we generally do not include the uncertainty on c_t to isolate the influence of the surveillance noise. Note that if estimates of c_t are available or derived from auxiliary data, these can be seamlessly integrated within our framework for more precise results.

Optimal model predictive control (MPC) of epidemics

The model predictive control algorithm we propose aims to curb disease spread, jointly minimising the risks and costs arising from infections and applied interventions. We outline our control framework in panel (a) of Fig. 1, which consists of the following elements: a plant (the controlled system) with observable states, state-transition probabilities, an agent with an action space defining possible control actions, and a reward function.

In our model, the plant is the population where the disease is spreading, while the output state monitored is the incidence (number of daily new infections) I_t. The state transition probabilities, i.e. the probability of transitioning to any I_{t+1} from any given I_t are not explicitly defined, but are implicitly determined by the Poisson distribution of the renewal model (see Eq. (1)). The control framework we use here largely overlaps with Markov decision processes (MDPs) [51]. However, the renewal model utilises both the immediate and past incidence. This is not exactly Markov but may be reconfigured into an MDP if higher dimensional state spaces are used [52, 53].

The agent in the context of an epidemic is the public health policy-maker i.e., the individual or group responsible for proposing or removing NPIs, while the action-space comprises the possible NPI choices. We consider 3 levels of interventions that we class as *no intervention*, *social distancing* and *full lockdown*. This broadly models stepped interventions which were common across the

COVID-19 pandemic. These include the three tier system that England used to enforce localised NPIs in 2020, the 4-level alert system applied by New Zealand and related policies taken by Italy, France, Canada and others [54, 55]. Our framework computes decisions based on the projected reward over a fixed time-horizon which incorporates the costs of possible actions in our decision space and their risks in terms of expected infections.

In the reward function, we account for costs arising from the economic impact of NPIs and the risks associated with high incidence. We consider a target incidence level that defines some manageable infection level and define the absolute error $I_{\text{err}} = |I_t - I_{\text{target}}|$. This target may relate to healthcare capacities e.g., setting a level of incidence such that the expected hospitalisations resulting from that incidence do not overwhelm healthcare resources. Although, studies rarely consider a target incidence level, our aim is to understand and characterise the intervention tradeoffs (e.g., timing choices) that can jointly limit expected infections and the costs of those interventions.

Setting $I_{\text{target}} = 0$ is analogous to an elimination target, which models the broad aims of pandemic policies employed by New Zealand [56] and China [57], for example. An $I_{\text{target}} > 0$ recognises that elimination is difficult, particularly in the face of infection reintroductions and so refocuses on stabilising healthcare burdens to sustainable levels that balance the supply and demand of health resources. Additionally, as we want to minimise the risk of large infection peaks and overshoots, our reward function also includes a penalty term ϕ_{over} that activates when $I_t > 1.5 I_{\text{target}}$ but is zero otherwise.

While regulating disease spread within the limits of healthcare capacities is of paramount importance, interventions that restrict mobility or close businesses and trade generate substantial economic and other costs. We model this with a term ϕ_t attached to every element of the action-space. There is no cost under no restrictions and the cost of full lockdown is assumed to be 15-times larger than that of limited social distancing. While some studies into COVID-19 NPIs suggest stringent interventions are 5-6 times more costly than more limited measures [7], our factor was chosen to more markedly distinguish between our two NPI tiers so that general qualitative insights could be better derived. Including all the above components, the reward function on day t is calculated as the negative quantity

$$r_t = -\delta |I_t - I_{\text{target}}| - \phi_t - \phi_{\text{over}}. \tag{4}$$

The agent's task is to choose the action which maximises the expected reward. However, there are practical limitations to decision-making. While we use daily incidence data to inform our epidemic model, we allow policy review to only occur every 7 or 14 days, i.e., the agent can only change control actions with this frequency. The time between policy updates is t_{rev} and reflects practical intervention constraints, e.g., both in terms of logistics and ensuring compliance, policymakers may not want to switch NPIs any faster than weekly. We also impose a practical constraint on reward optimisation by considering only finite time horizons for assessing the costs of any action. We denote this projection horizon t_{proj}. This models the fact that only short-term forecasts are known to be reliable for epidemic decision-making [58].

The agent calculates the expected reward for each action by simulating the epidemic with all possible control states until t_{proj} and taking the total temporally discounted reward

$$\rho = \sum_{s=0}^{t_{\text{proj}}} r_{t+s} \gamma^s, \tag{5}$$

where $\gamma < 1$ is the temporal discount factor. A higher γ means that the agent is more concerned about long-term rewards, whereas a smaller γ means that shorter term benefits are emphasised. Since the reward function is stochastic, the expected reward ρ is probabilistic. We therefore compute the expected total reward $\mathrm{E}(\rho)$ over an ensemble of simulations. This also allows us to factor in the intrinsic variability of the epidemic generating process (e.g., from the random times between infections).

If the projection horizon is longer than the policy review period, i.e., $t_{\text{pred}} > t_{\text{rew}}$, then, we can also propose sequences of actions over the projection horizon to be taken by the agent. We may then compute the expected reward for each action sequence then implement the first action of the sequence with the best projected reward. However, for the scenarios we consider, this approach increases computational complexity but does not improve performance. Consequently, for these longer projections, we maintain our original approach of only evaluating single possible actions and their consequences across the horizon. We collect all the key epidemic and control algorithm parameters in Table 1.

Parameter name		COVID-19	Ebola virus	Unit
Basic reproduction number	R_0	3.5	2.5	1
Mean generation time	$\alpha_\tau \beta_\tau$	6.5	15.0	days
Generation time variance	$\alpha_\tau \beta_\tau^2$	13.65	31.5	days^2
Infection incidence error gain	δ	0.00026	0.00065	$\frac{1}{\text{cases/day}}$
Projection window ($t_{\text{rev}} = 7$ days)	t_{pred}	12	12	days
Projection window ($t_{\text{rev}} = 14$ days)	t_{pred}	14	14	days
Cost of no intervention	ϕ_t	0	0	1
Cost of social distancing	ϕ_t	0.01	0.01	1
Cost of full lockdown	ϕ_t	0.15	0.15	1
Cost of large overshoots	ϕ_{over}	5.0	5.0	1
Cost discount factor	γ	0.95	0.95	1
Threshold for imposing full lockdown	C_{LD}	2500	3000	cases/day
Threshold for relaxing full lockdown	C_{relax}	1500	3500	cases/day
Lockdown length in cycle	t_{LD}	45	36	days
No intervention length in cycle	t_{relax}	9	21	days

Table 1: Parameters of the epidemic model and the control algorithms for COVID-19 and Ebola virus.

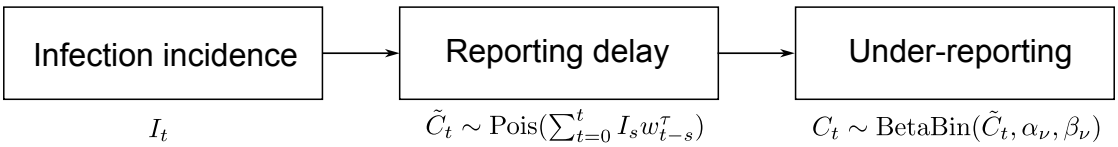

Figure 6: Models of realistic epidemic surveillance. The true infection incidence data I_t is first distorted by a probabilistic delay modelled by a convolution with w_{t-s}^τ, which are probabilities from a Gamma distribution. Under-ascertainment then occurs by downsampling these delayed cases $\tilde{C}_t$ using a Beta-binomial distribution. This yields the reported daily cases C_t, which is frequently used as a proxy for the unobservable I_t. In some simulations, we turn either reporting delay or under-reporting off. If there is no reporting delay, $\tilde{C}_t = I_t$ and similarly, if there is no under-reporting $C_t = \tilde{C}_t$.

Surveillance noise and uncertainty in incidence data

Ideally, the agent would make decisions about possible NPIs based on the infection incidence in the population. Unfortunately, infection data are rarely available and a proxy such as the incidence of confirmed cases or deaths is commonly used. We focus here on the daily incidence of cases C_t but note that other proxies have analogous descriptions [30, 31]. These proxies are commonly subject to practical surveillance imperfections, which we define via a stochastic reporting delay τ and a stochastic reporting rate ν. In our model of the surveillance imperfections, the true infection incidence data I_t is first distorted by delay, then we consider under-ascertainment of the delayed cases (see Fig. 6).

Reporting delay describes the lag between an infection and its proxy. For cases this includes latencies such as the time taken between infection and presenting symptoms or confirmation via testing. In our framework, we model the reporting delay for a single case using a Gamma distribution

$$\tau \sim \text{Gamma}(\alpha_\tau, \beta_\tau), \tag{6}$$

with shape and scale factors α_τ and β_τ, respectively. The mean reporting delay is then $\alpha_\tau \beta_\tau$ while the variance is $\alpha_\tau \beta_\tau^2$. To control the mean delay τ_{mean} and dispersion α directly we re-parametrise the distribution by the choice $\alpha_\tau = \alpha$ and $\beta_\tau = \tau_{\text{mean}}/\alpha$. The cases $\tilde{C}_t$ reported with delay on day t then result from a weighted sum of past incidence at day s and the probability that it takes $t - s$ time units before those infections are reported or confirmed as cases. This follows as

$$\tilde{C}_t \sim p^\tau(s) = \text{Pois}(\sum_{s=0}^{t} I_s w_{t-s}^\tau), \tag{7}$$

where the weight factors w_{t-s}^τ are derived from the Gamma distribution of the reporting delay, $w_{t-s}^\tau = p^\tau(t-s)/\sum_{k=0}^{t} p^\tau(k)$.

The reporting rate models incomplete or under-reporting, which captures the fact that proxies commonly represent only a fraction of infections. For example, asymptomatic and less severe infections are unlikely to appear as cases. This means that only a fraction ν of the delayed infection incidence $\tilde{C}_t$ is reported

$$C_t = \nu_t \tilde{C}_t, \tag{8}$$

with $\nu_t \in [0, 1]$. In our model, we assume that the number of reported cases follows a Beta-binomial distribution

$$C_t \sim \text{BetaBin}(\tilde{C}_t, \alpha_\nu, \beta_\nu). \tag{9}$$

Consequently, the expected number of reported cases is $\tilde{C}_t \alpha_\nu/(\alpha_\nu + \beta_\nu)$ while the variance is $\tilde{C}_t \alpha_\nu \beta_\nu (\alpha_\nu + \beta_\nu + \tilde{C}_t)/((\alpha_\nu + \beta_\nu)^2(\alpha_\nu + \beta_\nu + 1))$. We refer to the ratio of the expected reported cases and the true infection incidence as the mean reporting ratio $\nu := E[C_t/\tilde{C}_t]$ which is constant in time. To directly control the mean reporting ratio ν_{mean} and the dispersion a we choose $\alpha_n u = a$ and $\beta_n u = a(1 - \nu_{\text{mean}})/\nu_{\text{mean}}$.

In some cases, we investigate the isolated effect of reporting delay or under-reporting, which means that for these simulations the other source of surveillance imperfections is turned off. If there is no reporting delay, $\tilde{C}_t = I_t$ and similarly, if there is no under-reporting $C_t = \tilde{C}_t$. These stochastic delay [59] and under-reporting [60] models have been widely used to describe surveillance noise in the literature, as well as serve as the starting point for deconvolution and nowcasting methods that attempt to correct for these noise sources [31, 61, 62, 63].

Estimation of the reproduction number

When projecting likely infections (or proxies) over a horizon in our algorithm, we assume knowledge of the effect of NPIs on the reproduction number. This is captured by the coefficients c_t. However, we do not assume knowledge of the true basic reproduction number and so must estimate this quantity from past data.

We start by inferring the time-varying effective reproduction number by applying the formula [50]

$$R_t \approx \frac{\sum_{s=1}^{t_{\text{est}}} I_{t-s}}{\sum_{s=1}^{t_{\text{est}}} \Lambda_{t-s}}, \tag{10}$$

where Λ is the total infectiousness and is calculated as

$$\Lambda_{t-s} = \sum_{s=1}^{t_{\text{est}}} I_{t-s} w_s, \tag{11}$$

with weights w_s derived from the generation time distribution. Then we recover the basic reproduction number R_0 by factoring in the history of applied control measures

$$c_{\text{est}} = \frac{\sum_{t=0}^{t_{\text{est}}} c_{t-s} w_s}{\sum_{s=0}^{t_{\text{est}}} w_s} \tag{12}$$

and hence $R_{0\text{est}} = R_t/c_{\text{est}}$.

The quality of our estimates in Eqs.(10),(11) and (12) depends on the length of the estimation window t_{est}. Short windows are more sensitive to stochastic fluctuations in incidence, while long windows over-smooth estimates and delay projections [50, 64]. We apply $t_{\text{est}} = 5$ days, which appears to be a good compromise between the two extremes.

Code availability

The code generating the results presented here is available at `https://github.com/sandorberegi/Epidemic-control-with-noisy-real-time-data`.

Acknowledgements

SB and KVP acknowledge funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/X020258/1), funded by the UK Medical Research Council (MRC). This UK funded award is carried out in the frame of the Global Health EDCTP3 Joint Undertaking. The funders had no role in study design, data collection and analysis, decision to publish, or manuscript preparation. For the purpose of open access, the author has applied a 'Creative Commons Attribution' (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

SB: Formal Analysis, Investigation, Methodology, Software, Visualisation, Validation, Writing – original draft, Writing – review and editing. KP: Conceptualisation, Investigation, Methodology, Validation, Supervision, Funding Acquisition, Writing – review and editing.

Additional information

References

[1] F. Casella, "Can the COVID-19 epidemic be controlled on the basis of daily test reports?," *IEEE Control Systems Letters*, vol. 5, pp. 1079–1084, jul 2021.

[2] A. Mendez-Brito, C. El Bcheraoui, and F. Pozo-Martin, "Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19," *Journal of Infection*, vol. 83, pp. 281–293, Sept. 2021.

[3] N. Ferguson, D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. Van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, P. Winskill, C. Whittaker, C. Donnelly, S. Riley, and A. Ghani, "Report 9: Impact of non-pharmaceutical interventions (npis) to reduce COVID19 mortality and healthcare demand," 2020.

[4] J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, A. B. Stephenson, G. Leech, G. Altman, V. Mikulik, A. J. Norman, J. T. Monrad, T. Besiroglu, H. Ge, M. A. Hartwick, Y. W. Teh, L. Chindelevitch, Y. Gal, and J. Kulveit, "Inferring the effectiveness of government interventions against COVID-19," *Science*, vol. 371, Feb. 2021.

[5] T. P. B. Thu, P. N. H. Ngoc, N. M. Hai, and L. A. Tuan, "Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries," *Science of The Total Environment*, vol. 742, p. 140430, Nov. 2020.

[6] N. Huberts and J. Thijssen, "Optimal timing of interventions during an epidemic," *SSRN Electronic Journal*, 2020.

[7] D. J. Haw, G. Forchini, P. Doohan, P. Christen, M. Pianella, R. Johnson, S. Bajaj, A. B. Hogan, P. Winskill, M. Miraldo, P. J. White, A. C. Ghani, N. M. Ferguson, P. C. Smith, and K. D. Hauck, "Optimizing social and economic activity while containing SARS-CoV-2 transmission using DAEDALUS," *Nature Computational Science*, vol. 2, pp. 223–233, apr 2022.

[8] D. J. Haw, C. Morgenstern, G. Forchini, R. Johnson, P. Doohan, P. C. Smith, and K. D. Hauck, "Data needs for integrated economic-epidemiological models of pandemic mitigation policies," *Epidemics*, vol. 41, p. 100644, Dec. 2022.

[9] A. Lison, N. Banholzer, M. Sharma, S. Mindermann, H. J. T. Unwin, S. Mishra, T. Stadler, S. Bhatt, N. M. Ferguson, J. Brauner, and W. Vach, "Effectiveness assessment of non-pharmaceutical interventions: lessons learned from the COVID-19 pandemic," *The Lancet Public Health*, vol. 8, pp. e311–e317, Apr. 2023.

[10] E. Gubar, L. Policardo, E. J. Sánchez Carrera, and V. Taynitskiy, "On optimal lockdown policies while facing socioeconomic costs," *Annals of Operations Research*, June 2023.

[11] A. J. Meadows, B. Oppenheim, J. Guerrero, B. Ash, R. Badker, C. K. Lam, C. Pardee, C. Ngoon, P. T. Savage, V. Sridharan, N. K. Madhav, and N. Stephenson, "Infectious disease underreporting is predicted by country-level preparedness, politics, and pathogen severity," *Health Security*, vol. 20, pp. 331–338, Aug. 2022.

[12] K. V. Parag, "How to measure the controllability of an infectious disease?," Oct. 2023.

[13] C. M. Peak, L. M. Childs, Y. H. Grad, and C. O. Buckee, "Comparing nonpharmaceutical interventions for containing emerging epidemics," *Proceedings of the National Academy of Sciences*, vol. 114, pp. 4023–4028, Mar. 2017.

[14] G. Stépán and G. Haller, "Quasiperiodic oscillations in robot dynamics," *Nonlinear Dynamics*, vol. 8, pp. 513–528, Dec. 1995.

[15] G. Orosz, R. E. Wilson, and G. Stépán, "Traffic jams: dynamics and control," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, pp. 4455–4479, Oct. 2010.

[16] I. Vörös, G. Orosz, and D. Takács, "On the global dynamics of path-following control of automated passenger vehicles," *Nonlinear Dynamics*, vol. 111, pp. 8235–8252, Feb. 2023.

[17] Y. Kyrychko, K. Blyuss, P. Hövel, and E. Schöll, "Asymptotic properties of the spectrum of neutral delay differential equations," *Dynamical Systems*, vol. 24, pp. 361–372, Aug. 2009.

[18] H. T. Sykora, M. Sadeghpour, J. I. Ge, D. Bachrathy, and G. Orosz, "On the moment dynamics of stochastically delayed linear control systems," *International Journal of Robust and Nonlinear Control*, vol. 30, pp. 8074–8097, Oct. 2020.

[19] L.-S. Young, S. Ruschel, S. Yanchuk, and T. Pereira, "Consequences of delays and imperfect implementation of isolation in epidemic control," *Scientific Reports*, vol. 9, mar 2019.

[20] G. Albi, L. Pareschi, and M. Zanella, "Control with uncertain data of socially structured compartmental epidemic models," *Journal of Mathematical Biology*, vol. 82, May 2021.

[21] T. Britton and L. Leskelä, "Optimal intervention strategies for minimizing total incidence during an epidemic," *SIAM Journal on Applied Mathematics*, vol. 83, pp. 354–373, Mar. 2023.

[22] M. Bin, P. Y. K. Cheung, E. Crisostomi, P. Ferraro, H. Lhachemi, R. Murray-Smith, C. Myant, T. Parisini, R. Shorten, S. Stein, and L. Stone, "Post-lockdown abatement of COVID-19 by fast periodic switching," *PLOS Computational Biology*, vol. 17, p. e1008604, jan 2021.

[23] F. Di Lauro, I. Z. Kiss, and J. C. Miller, "Optimal timing of one-shot interventions for epidemic control," *PLOS Computational Biology*, vol. 17, p. e1008763, Mar. 2021.

[24] D. H. Morris, F. W. Rossine, J. B. Plotkin, and S. A. Levin, "Optimal, near-optimal, and robust epidemic control," *Communications Physics*, vol. 4, apr 2021.

[25] K. van Heusden, G. E. Stewart, S. P. Otto, and G. A. Dumont, "Effective pandemic policy design through feedback does not need accurate predictions," *PLOS Global Public Health*, vol. 3, p. e0000955, Feb. 2023.

[26] N. C. Grassly and C. Fraser, "Mathematical models of infectious disease transmission," *Nature Reviews Microbiology*, vol. 6, pp. 477–487, May 2008.

[27] M. Schwenzer, M. Ay, T. Bergs, and D. Abel, "Review on model predictive control: an engineering perspective," *The International Journal of Advanced Manufacturing Technology*, vol. 117, pp. 1327–1349, Aug. 2021.

[28] W. Uther, *Markov Decision Processes*, pp. 642–646. Springer US, 2011.

[29] N. G. Reich, D. A. T. Cummings, S. A. Lauer, M. Zorn, C. Robinson, A.-C. Nyquist, C. S. Price, M. Simberkoff, L. J. Radonovich, and T. M. Perl, "Triggering interventions for influenza: The ALERT algorithm," *Clinical Infectious Diseases*, vol. 60, pp. 499–504, Nov. 2014.

[30] K. V. Parag, C. A. Donnelly, and A. E. Zarebski, "Quantifying the information in noisy epidemic curves," *Nature Computational Science*, vol. 2, pp. 584–594, Sept. 2022.

[31] A. Azmon, C. Faes, and N. Hens, "On the estimation of the reproduction number based on misreported epidemic data," *Statistics in Medicine*, vol. 33, pp. 1176–1192, Oct. 2013.

[32] K. V. Parag and C. A. Donnelly, "Fundamental limits on inferring epidemic resurgence in real time using effective reproduction numbers," *PLOS Computational Biology*, vol. 18, p. e1010004, apr 2022.

[33] D. Chen, Y.-C. Lau, X.-K. Xu, L. Wang, Z. Du, T. K. Tsang, P. Wu, E. H. Y. Lau, J. Wallinga, B. J. Cowling, and S. T. Ali, "Inferring time-varying generation time, serial interval, and incubation period distributions for COVID-19," *Nature Communications*, vol. 13, Dec. 2022.

[34] M. Park, A. R. Cook, J. T. Lim, Y. Sun, and B. L. Dickens, "A systematic review of COVID-19 epidemiology based on current evidence," *Journal of Clinical Medicine*, vol. 9, p. 967, Mar. 2020.

[35] E. Shim, W. Choi, and Y. Song, "Clinical time delay distributions of COVID-19 in 2020–2022 in the Republic of Korea: Inferences from a nationwide database analysis," *Journal of Clinical Medicine*, vol. 11, p. 3269, June 2022.

[36] A. Tariq, Y. Lee, K. Roosa, S. Blumberg, P. Yan, S. Ma, and G. Chowell, "Real-time monitoring the transmission potential of COVID-19 in Singapore, March 2020," *BMC Medicine*, vol. 18, June 2020.

[37] A. R. Akhmetzhanov, H. Lee, S.-m. Jung, T. Kayano, B. Yuan, and H. Nishiura, "Analyzing and forecasting the Ebola incidence in North Kivu, the Democratic Republic of the Congo from 2018–19 in real time," *Epidemics*, vol. 27, pp. 123–131, June 2019.

[38] G. Pullano, L. D. Domenico, C. E. Sabbatini, E. Valdano, C. Turbelin, M. Debin, C. Guerrisi, C. Kengne-Kuetche, C. Souty, T. Hanslik, T. Blanchon, P.-Y. Boëlle, J. Figoni, S. Vaux, C. Campèse, S. Bernard-Stoecklin, and V. Colizza, "Underdetection of cases of COVID-19 in france threatens epidemic control," *Nature*, vol. 590, pp. 134–139, dec 2020.

[39] A. M. Carabelli, T. P. Peacock, L. G. Thorne, W. T. Harvey, J. Hughes, T. I. de Silva, S. J. Peacock, W. S. Barclay, T. I. de Silva, G. J. Towers, and D. L. Robertson, "SARS-CoV-2 variant biology: immune escape, transmission and fitness," *Nature Reviews Microbiology*, Jan. 2023.

[40] S. Contreras, J. P. Biron-Lattes, H. A. Villavicencio, D. Medina-Ortiz, N. Llanovarced-Kawles, and A. Olivera-Nappa, "Statistically-based methodology for revealing real contagion trends and correcting delay-induced errors in the assessment of COVID-19 pandemic," *Chaos, Solitons and amp; Fractals*, vol. 139, p. 110087, Oct. 2020.

[41] V. V. L. Albani, R. A. S. Albani, E. Massad, and J. P. Zubelli, "Nowcasting and forecasting COVID-19 waves: the recursive and stochastic nature of transmission," *Royal Society Open Science*, vol. 9, Aug. 2022.

[42] A. C. Miller, L. A. Hannah, J. Futoma, N. J. Foti, E. B. Fox, A. D'Amour, M. Sandler, R. A. Saurous, and J. A. Lewnard, "Statistical deconvolution for inference of infection time series," *Epidemiology*, vol. 33, pp. 470–479, May 2022.

[43] L. J. Beesley, D. Osthus, and S. Y. Del Valle, "Addressing delayed case reporting in infectious disease forecast modeling," *PLOS Computational Biology*, vol. 18, p. e1010115, June 2022.

[44] M. J. Keeling, L. Dyson, M. J. Tildesley, E. M. Hill, and S. Moore, "Comparison of the 2021 COVID-19 roadmap projections against public health data in England," *Nature Communications*, vol. 13, Aug. 2022.

[45] Z. Wang, M. A. Andrews, Z.-X. Wu, L. Wang, and C. T. Bauch, "Coupled disease–behavior dynamics on complex networks: A review," *Physics of Life Reviews*, vol. 15, pp. 1–29, Dec. 2015.

[46] B. Phillips and C. T. Bauch, "Early warning indicators of epidemics on a coupled behaviour-disease model with vaccine hesitance and incomplete data," *Journal of Dynamics and Games*, vol. 10, no. 1, pp. 49–86, 2023.

[47] A. Franzen and F. Wöhner, "Fatigue during the COVID-19 pandemic: Evidence of social distancing adherence from a panel study of young adults in Switzerland," *PLOS ONE*, vol. 16, p. e0261276, Dec. 2021.

[48] A. Petherick, R. Goldszmidt, E. B. Andrade, R. Furst, T. Hale, A. Pott, and A. Wood, "A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue," *Nature Human Behaviour*, vol. 5, pp. 1145–1160, Aug. 2021.

[49] M. D. Van Kerkhove, A. I. Bento, H. L. Mills, N. M. Ferguson, and C. A. Donnelly, "A review of epidemiological parameters from Ebola outbreaks to inform early public health decision-making," *Scientific Data*, vol. 2, May 2015.

[50] A. Cori, N. M. Ferguson, C. Fraser, and S. Cauchemez, "A new framework and software to estimate time-varying reproduction numbers during epidemics," *American Journal of Epidemiology*, vol. 178, pp. 1505–1512, sep 2013.

[51] M. Toussaint and A. Storkey, "Probabilistic inference for solving discrete and continuous state Markov Decision Processes," in *Proceedings of the 23rd international conference on Machine learning - ICML '06*, ICML '06, ACM Press, 2006.

[52] C. T. Baker, "Retarded differential equations," *Journal of Computational and Applied Mathematics*, vol. 125, pp. 309–335, Dec. 2000.

[53] S. Beregi, D. Takacs, and G. Stepan, "Bifurcation analysis of wheel shimmy with non-smooth effects and time delay in the tyre–ground contact," *Nonlinear Dynamics*, vol. 98, pp. 841–858, Aug. 2019.

[54] L. Smith, H. Potts, R. Amlohttps://doi.org/10.1016/S0140-6736(23)01368-5t, N. Fear, S. Michie, and G. Rubin, "Tiered restrictions for COVID-19 in England: knowledge, motivation and self-reported behaviour," *Public Health*, vol. 204, pp. 33–39, Mar. 2022.

[55] M. Campbell, L. Marek, J. Wiki, M. Hobbs, C. E. Sabel, J. McCarthy, and S. Kingham, "National movement patterns during the COVID-19 pandemic in New Zealand: the unexplored role of neighbourhood deprivation," *Journal of Epidemiology and Community Health*, vol. 75, pp. 903–905, Mar. 2021.

[56] S. Kung, T. Hills, N. Kearns, and R. Beasley, "New Zealand's COVID-19 elimination strategy and mortality patterns," *The Lancet*, vol. 402, pp. 1037–1038, Sept. 2023.

[57] J.-L. Tang and K. Abbasi, "What can the world learn from china's response to covid-19?," *BMJ*, p. n2806, Dec. 2021.

[58] B. Prasse, M. A. Achterberg, and P. Van Mieghem, "Accuracy of predicting epidemic outbreaks," *Physical Review E*, vol. 105, p. 014302, Jan. 2022.

[59] K. M. Gostic, L. McGough, E. B. Baskerville, S. Abbott, K. Joshi, C. Tedijanto, R. Kahn, R. Niehus, J. A. Hay, P. M. De Salazar, J. Hellewell, S. Meakin, J. D. Munday, N. I. Bosse, K. Sherrat, R. N. Thompson, L. F. White, J. S. Huisman, J. Scire, S. Bonhoeffer, T. Stadler, J. Wallinga, S. Funk, M. Lipsitch, and S. Cobey, "Practical considerations for measuring the effective reproductive number, Rt," *PLOS Computational Biology*, vol. 17, p. e1009679, Dec. 2021.

[60] K. M. Gamado, G. Streftaris, and S. Zachary, "Modelling under-reporting in epidemics," *Journal of Mathematical Biology*, vol. 69, pp. 737–765, Aug. 2013.

[61] L. S. Bastos, T. Economou, M. F. C. Gomes, D. A. M. Villela, F. C. Coelho, O. G. Cruz, O. Stoner, T. Bailey, and C. T. Codeço, "A modelling approach for correcting reporting delays in disease surveillance data," *Statistics in Medicine*, vol. 38, pp. 4363–4377, July 2019.

[62] S. F. McGough, M. A. Johansson, M. Lipsitch, and N. A. Menzies, "Nowcasting by Bayesian smoothing: A flexible, generalizable model for real-time epidemic tracking," *PLOS Computational Biology*, vol. 16, p. e1007735, Apr. 2020.

[63] E. Goldstein, J. Dushoff, J. Ma, J. B. Plotkin, D. J. D. Earn, and M. Lipsitch, "Reconstructing influenza incidence by deconvolution of daily mortality time series," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 21825–21829, Dec. 2009.

[64] K. V. Parag and C. A. Donnelly, "Using information theory to optimise epidemic models for real-time prediction and estimation," *PLOS Computational Biology*, vol. 16, p. e1007990, July 2020.

Supplement

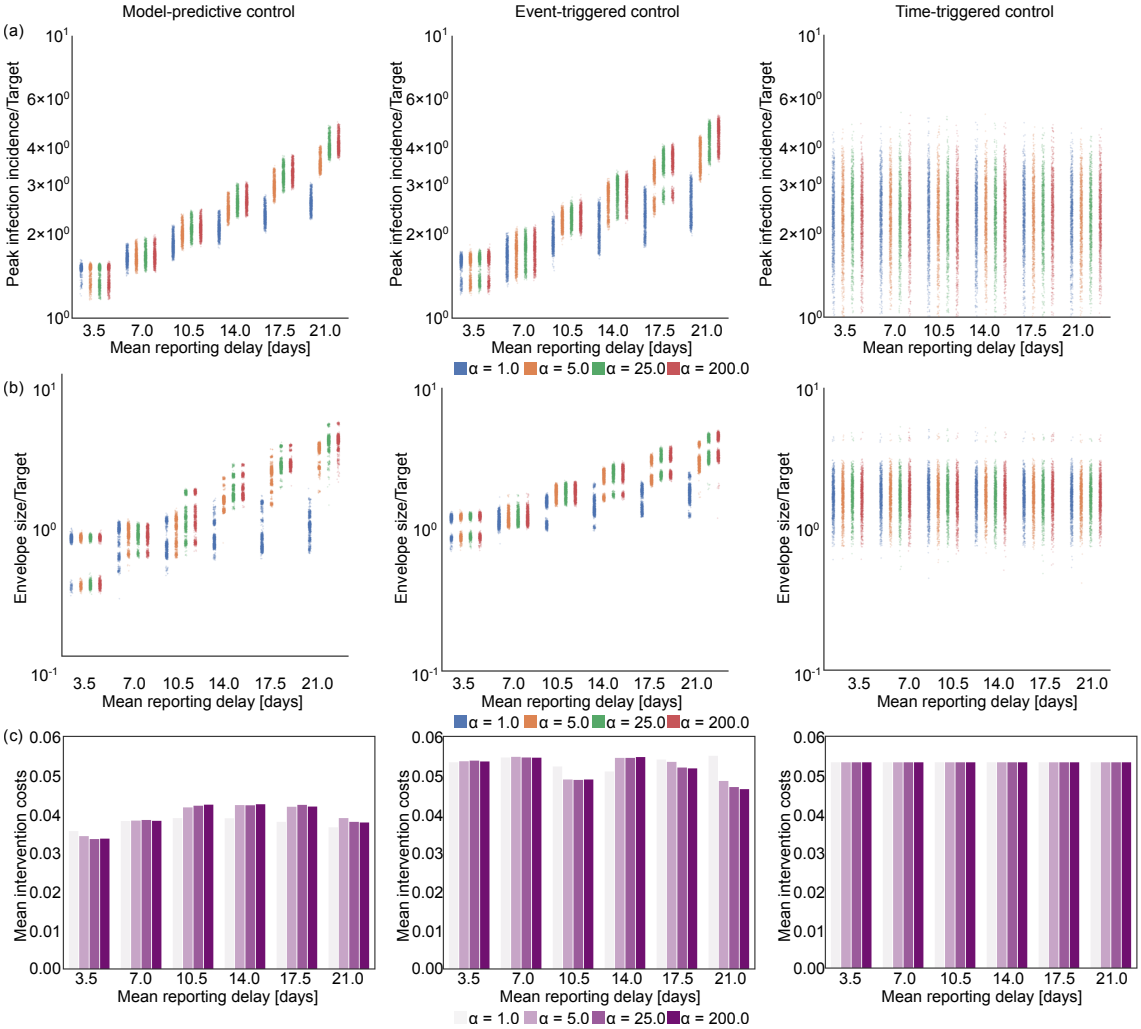

Figure 7: The impact of increasing reporting delay on optimal control. The left panels show results for our MPC algorithm, while the middle and right panels respectively present equivalent outputs under event-triggered (lockdown and relaxation thresholds of $C_{\mathrm{LD}} = 3000$ new cases/day, and $C_{\mathrm{relax}} = 3500$ new cases/day) and time-triggered (cycle of 36 days lockdown, 21 days of no restrictions, starting on day 143) strategies. Row (a) shows scatterplots for peak incidence and row (b) illustrates the steady-state envelope size relative to the target incidence across different time delay distributions. The horizontal axis represents different mean reporting delays with colours depicting different dispersion levels. Larger values of the parameter α indicate more deterministic delays. Row (c) shows the mean intervention costs for each ensemble of 1000 epidemics, simulated under estimated parameters from Ebola virus disease.

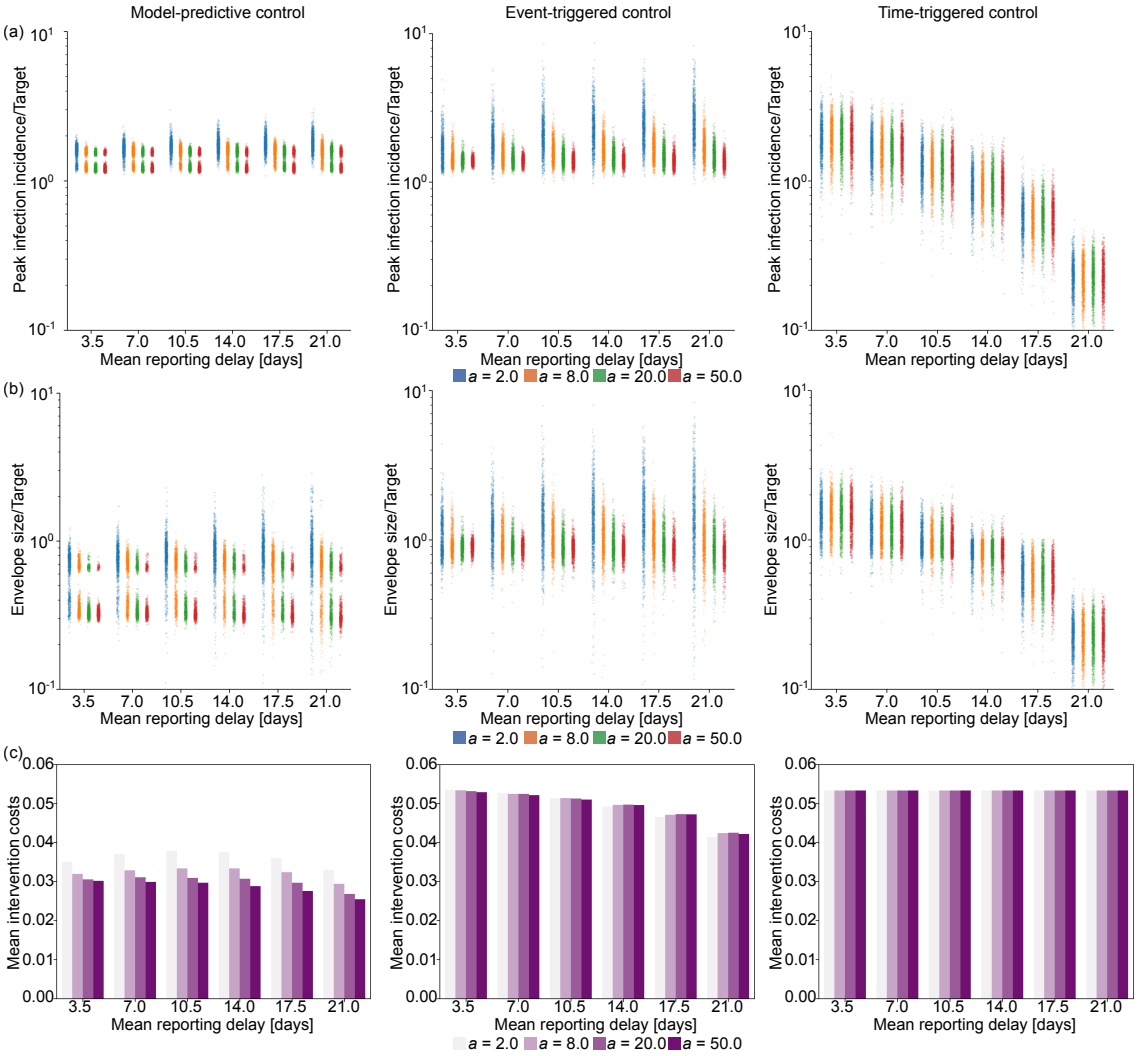

Figure 8: The impact of increasing under-reporting on optimal control. The left panels show results for our MPC algorithm, while the middle and right panels respectively present equivalent outputs under event-triggered (lockdown and relaxation thresholds of $C_{\mathrm{LD}} = 3000$ new cases/day, and $C_{\mathrm{relax}} = 3500$ new cases/day) and time-triggered (cycle of 36 days lockdown, 21 days of no restrictions, starting on day 143) strategies. Row (a) shows scatterplots for peak incidence, row (b) illustrates the steady-state envelope size relative to the target incidence for different reporting rate distributions. The horizontal axis represents different mean reporting ratios with colours depicting different dispersion levels. Larger values of the parameter a belong to more deterministic case reporting distributions (i.e., constant reporting). Row (c) shows the mean intervention costs for each ensemble of 1000 epidemics, simulated under estimated parameters from Ebola virus disease.

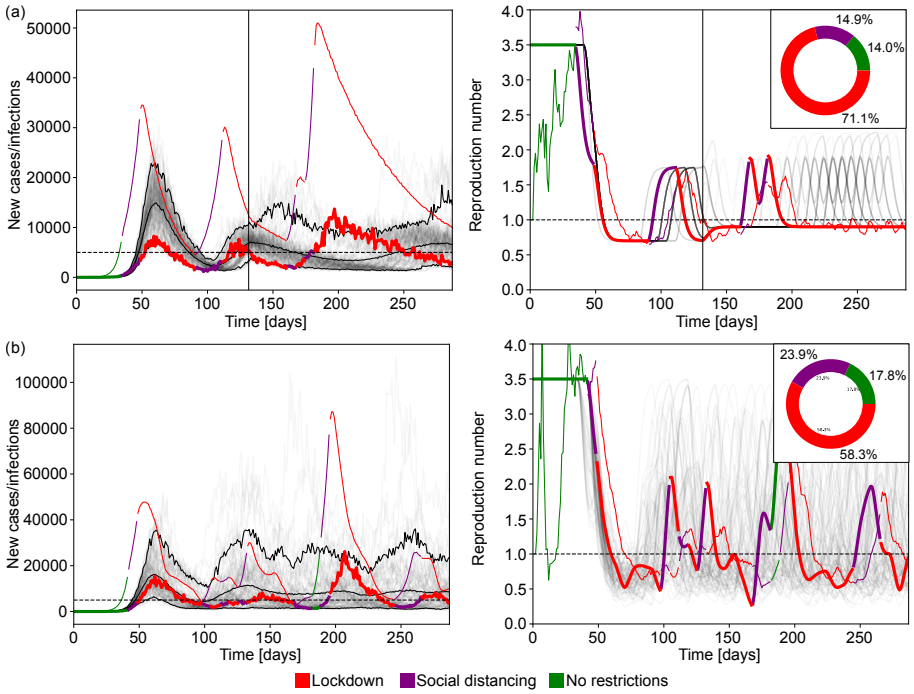

Figure 9: Optimal control for realistic simulated epidemics with uncertainty in the estimated reproduction number. Row (a) shows epidemics simulated under COVID-like generation time where the basic reproduction number is changed from 3.5 to 4.5 on day 130 (black vertical line). Row (b) represents epidemics simulated under COVID-like generation time and a fixed basic reproduction number with uncertainty in the effect of NPIs on reduction in disease transmissions. Instead of fixed reduction factors, the factor c_t altering the basic reproduction number as $R_{0t} = R_0 c_t$ is sampled form a Beta distribution. The left column show reported cases from ensembles of 100 simulations with mean delay of 10.5 days, delay dispersion $\alpha = 5.0$, mean reporting rate 0.25, $a = 8.0$. The faded curves show different individual realisations of the epidemic with the three black curves marking the 5% and 95% percentiles of the ensemble and the mean reported cases. The horizontal dashed line shows the incidence target. The highlighted thick curve of reported cases is coloured based on the NPI implemented on a given day. The coloured thin curve indicates the true incidence corresponding to that highlighted simulation. The right column shows similar diagrams for the effective reproduction number. The faded grey curves and the highlighted curve represents the effective reproduction number estimated from true incidence whereas the thin curve indicates the estimated value from the reported cases in the highlighted realisation of the epidemic. The inset pie charts in the right column indicate the ratio of days spent under a given NPI across the full simulation ensemble.