OPEN

Optimal national prioritization policies for hospital care during the SARS-CoV-2 pandemic

Josh C. D'Aeth^{1,2,9}, Shubhechya Ghosal^{1,2,9}, Fiona Grimm^{1,2,9}, David Haw^{1,2,9}, Esma Koca^{1,2,9}, Krystal Lau^{1,2,9}, Stefano Moret^{1,2,9}, Dheeya Rizmie^{1,2,9}, Sarah R. Deeny⁴, Pablo N. Perez-Guzman^{1,2}, Neil Ferguson^{1,2}, Katharina Hauck^{1,2}, Peter C. Smith^{5,6,7}, Giovanni Forchini^{1,2,8}, Wolfram Wiesemann³ and Marisa Miraldo^{1,2,9}✉

In response to unprecedented surges in the demand for hospital care during the SARS-CoV-2 pandemic, health systems have prioritized patients with COVID-19 to life-saving hospital care to the detriment of other patients. In contrast to these ad hoc policies, we develop a linear programming framework to optimally schedule elective procedures and allocate hospital beds among all planned and emergency patients to minimize years of life lost. Leveraging a large dataset of administrative patient medical records, we apply our framework to the National Health Service in England and show that an extra 50,750–5,891,608 years of life can be gained compared with prioritization policies that reflect those implemented during the pandemic. Notable health gains are observed for neoplasms, diseases of the digestive system, and injuries and poisoning. Our open-source framework provides a computationally efficient approximation of a large-scale discrete optimization problem that can be applied globally to support national-level care prioritization policies.

Health systems worldwide are struggling to provide hospital treatment during the surges in demand for emergency care caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic, despite efforts to increase capacity^{1–4}. To manage demand, many countries have prioritized patients with coronavirus disease 2019 (COVID-19), canceled elective (that is, planned) procedures^{5–7} and rationed access to life-saving care for all patients^{3,4}. For example, policies in Italy, in March 2020, involved prioritizing intensive care to patients with COVID-19 above 70 years who previously had no more than one admission per year for a chronic illness (for example, exacerbated chronic obstructive pulmonary disease, advanced neoplasms and congestive heart failure)⁸. In England, the cancellation of non-urgent elective surgeries after 17 March 2020 was combined with the prioritization to critical care (CC) of those with high capacity to benefit as signaled by a low frailty score^{5,6}. Even though several vaccines are now available, as the pandemic progresses, hospitals are still challenged by capacity constraints to schedule elective surgeries^{3,4,9}.

Policies that prioritize patients with COVID-19 over other patients with greater capacity to benefit may result in immediate or delayed deaths and morbidity among both non-COVID-19 patients and patients with COVID-19 and increase the financial burden on health systems as delays in planned treatments may accelerate disease progression and the need for more costly interventions later^{10,11}. Also, when implemented, these policies generate a backlog of non-COVID-19 patients in need of care^{12,13} that require prioritization rules that differ from pre-pandemic ones in order to be managed, since heterogeneity in disease progression over the postponement period might change their relative priority when compared with

other patients. In England, the National Health Service (NHS) Confederation has projected waiting lists to reach 9.8 million by the end of 2021¹⁴, highlighting how essential it is to identify ways to prioritize care and prevent hospitals from being overwhelmed by the various constraints posed by the pandemic¹⁵.

The optimal scheduling of hospital care has been studied albeit focusing on specific diseases^{16,17}, types of care^{18–20} or individual hospitals²¹. These studies deploy methodologies that allow for a faithful modeling of the operations of an individual hospital or a specific care setting and thus enable prioritization of subgroups of patients, but they do not scale to the health system of an entire country and all patients in need of care, which is the focus of our paper.

In many health systems worldwide, hospital scheduling relies on a set of prioritization rules, which seldom look at the entirety of the health system and therefore lead to suboptimal decisions. The need to mitigate this gap in evidence has been explicitly raised during the pandemic. In England, the Nuffield Council on Bioethics issued a national statement on the need for clearer and “nationally developed and coordinated guidance on how decisions about the allocation of constrained resources should be made,”²² highlighting also the need for prioritization to consider all disease groups and their capacity to benefit. Therefore, in contrast to the approach followed in the literature, we propose a framework to optimally schedule elective care and allocate hospital capacity to elective and emergency patients in general and acute (G&A) and CC across all disease areas while accounting for national capacity and needs.

Our approach relies on a linear programming (LP) formulation of the problem, which we present in abridged form in Methods and in its entirety in the Supplementary Information.

¹MRC Centre for Global Infectious Disease Analysis and WHO Collaborating Centre for Infectious Disease Modelling, School of Public Health, Imperial College London, London, UK. ²Abdul Latif Jameel Institute for Disease and Emergency Analytics (J-IDEA), School of Public Health, Imperial College London, London, UK. ³Department of Analytics, Marketing and Operations, Imperial College Business School, Imperial College London, London, UK. ⁴The Health Foundation, London, UK. ⁵Department of Economics and Public Policy, Imperial College Business School, Imperial College London, London, UK. ⁶Centre for Health Economics and Policy Innovation, Imperial College Business School, Imperial College London, London, UK. ⁷Centre for Health Economics, University of York, York, UK. ⁸Umeå School of Business, Economics and Statistics, Umeå University, Umeå, Sweden. ⁹These authors contributed equally: Josh C. D'Aeth, Shubhechya Ghosal, Fiona Grimm, David Haw, Esma Koca, Krystal Lau, Stefano Moret, Dheeya Rizmie. ✉e-mail: m.miraldo@imperial.ac.uk

Fig. 1 | Input-output overview of the LP model formulation for the OS. Inputs of the model are (1) the initial situation (at $t = 0$) in terms of the available resources (Ξ) and the current allocation of patients (waiting versus in-hospital patients, in CC versus G&A, and so on). Patients are divided into different patient groups and subdivided on the basis of severity. For each subgroup, we provide as inputs (2) their resource requirements (Δ) as well as transition matrices (Π) representing the probabilities of endogenous transfers of patients between severity groups (for example, patients needing emergency care while waiting for elective care, or patients in G&A requiring CC). For $t > 0$, based on the scenarios we are investigating (for example, lockdown), we observe (3) new exogenous inflows of patients (Φ). During each week, the model optimizes the allocation of patients, that is, how many patients of each group to admit to hospital (z_t, z'_t) as well as the in-hospital transfers of patients.

We apply our model to the NHS in England between 2 March 2020 and 1 March 2021, with the aim of minimizing years of life lost (YLL) under alternative scenarios. These consider capacity constraints, demand for emergency care and projected COVID-19 emergencies under a set of pandemic trajectories that reflect varying stringency of control strategies, projected with a susceptible–exposed–infected–recovered (SEIR) model of SARS-CoV-2 transmission. We compare the outcomes of our optimized schedules to those of a set of simulated prioritization policies that reflect those implemented by various health systems during the pandemic, including: (1) postponement of elective procedures; (2) prioritization to CC based on frailty; and (3) re-scheduling patients to elective care using pre-pandemic prioritization rules.

Results

Model overview. Optimized schedules (OS) are formulated with a deterministic LP model that optimally schedules the admission of patients to hospital across all diseases and nationally. Specifically, the key decision variables of the model are which patients to admit and when (admission scheduling), and which patients to allocate to CC in case of capacity shortages. In the OS, emergency patients are always admitted if capacity is available, while elective admissions are scheduled weekly over the 52-week horizon with the objective of minimizing YLL. Figure 1 offers a conceptual input–output overview of the LP formulation.

When demand exceeds capacity, OS ration care. First, patients in need of elective care remain on the waiting list. Second, patients requiring CC are treated in G&A (with different transition probabilities) until CC capacity becomes available. Third, during the peaks of the pandemic and under some scenarios, emergency capacity is not sufficient to accommodate all patients. Given that we cannot rely on pre-pandemic data to estimate the implications of

denied emergency care, our analysis considers two boundary cases. Our LP model conservatively assumes that all patients who are denied emergency admission die (upper-bound case). Alternatively, we use the solution of our LP model but re-evaluate the YLL under the assumption that the denied patients' probabilities and outcomes remain as if they would be cared for in hospital (lower-bound case). By modeling both extreme cases, we fully characterize the spectrum of reasonable outcomes.

We compare our OS with standard policies (SP) that simulate prioritization policies implemented in England during the pandemic. SP manage surges in demand by postponing elective admissions. Patients with postponed procedures are kept waiting and admitted once there is capacity according to a first-in-first-out rule, that is, following an order that preserves their assigned priority when care was originally scheduled. Different disease groups are admitted in equal proportions. Under SP, CC is prioritized for non-frail patients. All other assumptions are as in OS.

We model four SP informed by elective care postponements and actual admission policies in England. SP1 and SP3 assume that elective procedures were canceled between 17 March and 29 April 2020 (weeks 3–8), as actually occurred⁷. SP2 and SP4 additionally trigger cancellations anytime over the intervention horizon when the projected COVID-19 incidence exceeds specific trigger points. SP1 and SP2 assume 100%, SP3 and SP4 75% elective cancellations.

The incremental cost-effectiveness ratio (ICER, a standard metric used by policymakers to decide on interventions to implement in health systems) of OS versus SP is calculated across all disease groups and ages as (where YLG is years of life gained):

$$\text{ICER} = \frac{\Delta \text{Costs}}{\Delta \text{YLG}} = \frac{\text{Cost}_{\text{OS}} - \text{Cost}_{\text{SP}}}{\text{YLL}_{\text{SP}} - \text{YLL}_{\text{OS}}}$$

Model data. To forecast non-COVID-19 elective and emergency needs (CC and G&A), we use administrative patient medical records (Hospital Episode Statistics²³ (HES)) from all public acute hospitals in England between January 2015 and February 2020. To model care pathways of hospital patients with COVID-19, we use medical records from 614 patients admitted with SARS-CoV-2 infection to Imperial College Healthcare NHS hospitals between 25 February 2020 and 5 April 2020²⁴. YLL are calculated using standard life tables²⁵. Patients are individually costed using the National Cost Collection dataset from 2015 to 2019²⁶ matched to HES data at Health Resource Group (HRG) level. Staff numbers are obtained from the NHS Electronic Staff Records for 2020. G&A beds are calculated using the March 2020 extract of the Quarterly Bed Availability and Occupancy Dataset²⁷. CC beds are obtained from the Critical Care Monthly Situation Reports dataset for February 2020²⁸. Monthly emergency admissions are obtained from the A&E Attendances and Emergency Admissions dataset from NHS England Statistics for March to June 2020²⁹.

Scenario analysis. For both OS and SP as well as under both the upper- and the lower-bound assumptions, we account for uncertainty by considering alternative scenarios based on capacity constraints, inflows of emergency admissions and pandemic trajectories. For the capacity constraints, we consider that capacity either remains at pre-pandemic levels or is expanded to reflect hospital interventions introduced to increase total capacity by 16,500 beds (for example, field hospitals) and 38,462 staff³⁰ (for example, recruitment of retired and student medical staff), see Supplementary Section 3.2. For the inflows of emergency admissions, we consider that emergency admissions either remain as forecasted or that there is a reduction in the forecasted number of patients in need of emergency care (to reflect behavioral changes due to the pandemic and its mitigation strategies as well as potentially increased mortality at home). We reduce our forecasted emergency needs by 34%, using A&E attendance data to estimate the proportion of the reduction in emergency admissions throughout the pandemic (Supplementary Section 3.2). For the pandemic trajectories, and to capture a range of scenarios characterizing the severity of the pandemic and the pressure on hospital capacity, we consider the following reproduction numbers (R_t): $R_t = 1.2$, which yields a second-wave peak in hospital occupancy similar to the peak of the first wave; and $R_t = 1.1$, which generates a notably milder second wave of the pandemic, and a peak hospital occupancy akin to seasonal flu. These are combined with two alternative control interventions that impose lockdowns on 1 December 2020 (early LD) and 1 January 2021 (late LD).

The combination of these different assumptions leads to the following scenarios: baseline early-LD and late-LD scenarios assume an effective reproduction number of $R_t = 1.1$, pre-pandemic staff and bed capacity, and no reduction in emergencies. Best-case early- and late-LD scenarios assume $R_t = 1.1$ plus a 34% reduction in emergencies, and capacity expanded. Worst-case early- and late-LD scenarios assume $R_t = 1.2$, no reduction in emergencies and pre-pandemic capacity (Supplementary Section 8).

Scenario results. Except where flagged, YLL are reported as the minimum (which occurs in either the best-case early-LD or the best-case late-LD scenarios) and the maximum (which occurs in either the worst-case early-LD or the worst-case late-LD scenarios) observed YLL differences across all scenarios considering both the upper- and lower-bound results. Values reported refer to the comparison of OS with SP1. Comparisons with SP2–4 are in Supplementary Fig. 6.

YLL are greater under SP than OS in total (50,750–5,891,608 YLL across the scenarios) and for most disease groups and scenarios (Fig. 2). In per-capita terms and across all diseases, the largest gains accrue in the worst-case LD scenario with the OS offering up

to 1.58 and 0.82 additional years of life for patients aged <25 and aged 25–64, respectively. The International Statistical Classification of Diseases and Related Health Problems (ICD) groups with the largest gains are neoplasms (C00–D48: 18,111–144,719 YLG; with the highest observed per-capita gains observed for young patients amounting to 1.1 YLG), digestive system diseases (K00–K93: 3,935–606,016 YLG; highest observed per-capita gains for young patients amounting to 1.2 YLG), and injuries and poisoning (S00–T98: 4,818–742,851 YLG; highest observed per-capita gains for young patients amounting to 1.6 YLG). Large differences are also observed for diseases of the circulatory system (I00–I99: –253–368,797 YLG; highest observed per-capita gains for young patients amounting to 2.2 YLG) and for respiratory diseases (J00–J99: –385–1,131,860 YLG; highest observed per-capita gains for young patients amounting to 1.9 YLG), see Fig. 2. OS prioritize these patients (to emergency care and to CC) over patients with COVID-19, resulting in lower YLL than SP. The highest per-capita gains of OS over SP occur in the worst-case late-LD scenario for young (<25) patients with COVID-19 (3.8 YLG) and for diseases of the circulatory system (2.2 YLG). The OS penalize the elderly patients with COVID-19 with 3.4 YLL compared with SP.

The reduction in YLL achieved by OS is greater in scenarios where capacity constraints are more restrictive (worst-case upper bound), suggesting that an optimal prioritization of electives is increasingly beneficial as resources become scarce. When capacity enables accommodating all emergencies or there is scope to invest in extra capacity for emergencies beyond existing levels (lower bound), there is less difference between OS and SP because all patients can receive care. The notable health gains of OS do not come at increased costs in baseline and best-case scenarios (Fig. 3 and Table 1). For scenarios in which OS is costlier than SP, small extra spending is required to save many lives, with the extra costs being associated with an increased number of elective admissions and shifting from low-priority to costlier high-priority patients. In the worst-case scenarios, OS are cost effective for thresholds between £57 and £1,070 per YLL (Fig. 3 and Table 1).

OS accommodate more elective admissions than SP since there is no blanket postponement of elective procedures, and admission scheduling is determined by the patient's probability of survival and likelihood of needing emergency care while waiting for elective care. When capacity constraints are less restrictive (baseline and best-case early LD/late LD, worst-case early LD), OS exhibit fewer non-COVID-19 patients requiring emergency care than SP (Table 1). The ICDs for which the difference in elective (emergency) admissions is the highest (lowest) are digestive system diseases (K00–K93), neoplasms (C00–D48), genitourinary system disease (N00–N99), and injuries and poisoning (S00–T98) (Fig. 4 for SP1 and Supplementary Fig. 7 for SP2–4).

In the worst-case scenarios, capacity constraints are more stringent than in the baseline and best-case scenarios, and more patients require emergency care while waiting for elective care under both OS and SP. Due to resource constraints, differences in emergency admissions between OS and SP are smaller in the worst-case late-LD scenario (Table 1). Notice that there are no denials in emergency admissions in the best-case scenarios for either model. However, this is not the case in any other scenario, where emergency patients are denied admission due to capacity constraints. The number of emergency admission denials is higher for SP than OS (Table 1).

At baseline, OS deny CC to few patients (2.2% and 2.6% over the 52 weeks under early LD and late LD, respectively), most of which (83% and 85%) are patients with COVID-19 aged 65+ (Supplementary Fig. 9). In the worst-case scenarios, the share of patients denied CC increases to 4.8% and 6.4% under early LD and late LD, with 92% and 83% of patients denied CC stemming from COVID-19 aged 65+ (Fig. 5). Therefore, OS show that YLL are minimized when non-COVID-19 patients are prioritized

Fig. 2 | Comparison of SP1 and OS for YLL under various scenarios. Each dot represents the log of the difference in years of life lost (YLL) for all admissions under SP1 and the OS ($YLL_{OS} - YLL_{SP}$) over the 52-week planning horizon for each ICD-10 diagnosis code. Yellow dots represent the upper-bound differences. Dark purple dots represent the lower-bound differences. The vertical black dashed lines represent no difference (of zero YLL). Supplementary Fig. 6 presents this information for SP2–4. ICD-10 codes included are A00–B99 (certain infectious and parasitic diseases), C00–D48 (neoplasms), E00–E90 (endocrine, nutritional and metabolic diseases), F00–F99 (mental and behavioral disorders), G00–G99 (diseases of the nervous system), H00–H59 (diseases of the eye and adnexa), I00–I99 (diseases of the circulatory system), J00–J99 (diseases of the respiratory system), K00–K93 (diseases of the digestive system), L00–L99 (diseases of the skin and subcutaneous tissue), M00–M99 (diseases of the musculoskeletal system and connective tissue), N00–N99 (diseases of the genitourinary system), O00–O99 (pregnancy, childbirth and the puerperium), R00–R99 (symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified), S00–T98 (injury, poisoning and certain other consequences of external causes) and V01–Y98 (external causes of morbidity and mortality). The elective bundle is composed of ICD codes A00–B99, E00–E90, F00–F99, H60–H95 (diseases of the ear and mastoid process), O00–O99, P00–P96 (certain conditions originating in the perinatal period), and Q00–Q99 (congenital malformations, deformations and chromosomal abnormalities). The emergency bundle is composed of ICD codes D50–D89 (disease of blood, blood-forming organs, and certain disorders involving the immune mechanism), H00–H59, H60–H95, P00–P96, Q00–Q99 and Z00–Z99 (factors influencing health status and contact with health services).

for CC over patients with COVID-19. In the worst-case late-LD OS, patients with neoplasms and diseases of the circulatory system are prioritized over patients with COVID-19 aged 65+ (see Supplementary Fig. 10 for G&A bed utilization). In the best-case scenarios, OS admit to CC all patients requiring it, while SP deny CC to patients during the weeks following the cancellation of elective procedures (Table 1).

Discussion

In England, fears over premature mortality and morbidity associated with the postponement of elective procedures prompted policy

discussions on how to accommodate patients needing elective care⁴. NHS England directed hospitals to sustain elective care over the winter 2020–2021. This directive is proving difficult to meet with demands from patients with COVID-19 surging^{31,32}. Our findings are timely for the NHS as they have the potential to support re-scheduling delayed elective procedures while coping with further increases in patients with COVID-19.

While there is uncertainty about our estimates, reflected in the variability of the orders of magnitude of the reported gains of the OS over SP across the different scenarios, the scenarios and respective parameters reflect a broad spectrum of stringency of capacity

Fig. 3 | Incremental cost-effectiveness ratios/cost-effectiveness plane. Illustration of the incremental benefit (measured as years of life gained, YLG) and the incremental cost (in GBP, £) across two panels outlining incremental cost effectiveness for upper- and lower-bound estimates. These are calculated as $\text{Incremental cost} = \text{Cost}_{\text{OS}} - \text{Cost}_{\text{SP}}$ and $\text{Incremental benefit} = \text{YLL}_{\text{SP}} - \text{YLL}_{\text{OS}}$. The willingness-to-pay (WTP) threshold is drawn as a dashed red line with a value of £20,000 per YLG. Values below the WTP in the top-right quadrant imply that the OS are cost effective when compared with SP. Values in the bottom-right quadrant indicate that the OS are cheaper and are associated with higher YLG than the SP. Gray panels represent an enlarged view of their respective white panels. Different policy comparisons are illustrated by color: purple compares the OS with SP1, blue compares OS with SP2, green compares OS with SP3, and yellow compares OS with SP4. Different symbols represent different scenarios: dots illustrate the baseline early LD, triangles illustrate the baseline late LD, squares illustrate the best-case early LD, crosses illustrate the best-case late LD, crossed boxes illustrate the worst-case early LD and stars represent the worst-case late LD.

constraints. As such, gains of OS over SP are characterized across plausible states of the world. Importantly, our results highlight some high-level commonalities across all scenarios that are useful to inform the formulation of national guidance akin to the guidance developed by the National Institute for Excellence Health and

Care Excellence (NICE). Our findings outline life-saving key prioritization principles that can be embedded in national policies: (1) consider the relative capacity to benefit of patients with different diseases and ages; (2) postpone elective treatments for which disease progression is mild and that have lower chances of being

Table 1 | Health economic metrics and patient flows comparing OS versus SP1

Scenario name		Health economic metrics				
		Total cost ^a (£ millions)	Total YLL ^a (×1,000)	ICER calculations (OS versus SP1)		
				Incremental YLG ^a (×1,000)	Incremental costs ^a (£ millions)	ICER ^a (£ per YLG)
OS	Baseline early LD	[23,173; 23,180]	[5,113; 5,129]			
	Baseline late LD	[23,260; 23,268]	[5,350; 5,365]			
	Best-case early LD	[20,458; 20,458]	[4,313; 4,313]			
	Best-case late LD	[20,525; 20,525]	[4,548; 4,548]			
	Worst-case early LD	[23,197; 23,204]	[6,063; 6,078]			
	Worst-case late LD	[22,876; 23,151]	[7,266; 7803]			
SP1	Baseline early LD	[23,586; 23,611]	[5,432; 5,687]	[319; 559]	[−430; −413]	OS dominates
	Baseline late LD	[23,377; 23,401]	[5,669; 5,918]	[320;553]	[−133; −116]	OS dominates
	Best-case early LD	[20,506; 20,506]	[4,375; 4,375]	[61;61]	[−48; −48]	OS dominates
	Best-case late LD	[20,573; 20,573]	[4,599; 4,599]	[51; 51]	[−48; −48]	OS dominates
	Worst-case early LD	[22,922; 22,974]	[6,581; 7,130]	[518; 1,052]	[230; 275]	[445; 262]
	Worst-case late LD	[22,543; 23,097]	[7,864; 13,694]	[598; 5,892]	[54; 333]	[91; 57]
		Patient flows (×1,000)				
Scenario name		COVID emergency admissions	Non-COVID emergency admissions	Total elective admissions	Total admission denials	
OS	Baseline early LD	179	6,554	2,205	2	
	Baseline late LD	200	6,540	2,229	2	
	Best-case early LD	182	4,294	3,868	0	
	Best-case late LD	202	4,294	3,869	0	
	Worst-case early LD	249	6,533	2,156	2	
	Worst-case late LD	311	6,487	2,021	74	
SP1	Baseline early LD	181	6,635	2,154	9	
	Baseline late LD	201	6,638	2,014	8	
	Best-case early LD	182	4,318	3,850	0	
	Best-case late LD	202	4,318	3,850	0	
	Worst-case early LD	251	6,643	1,702	19	
	Worst-case late LD	329	6,487	1,600	196	

^aBrackets represent [lower bound; upper bound].

admitted to emergency care; (3) prioritize access to emergency care and critical care based on capacity to benefit, rather than by default prioritizing one disease (for example, COVID-19) over other patients (for example, non-COVID-19)². Where data are not available, these general principles can be used by policymakers globally to rapidly formulate national guidance in the future when national health services experience severe demand shocks such as those experienced during the pandemic. Our model is open source, and it is solved within minutes on a standard computer^{33,34}.

While our objective is to inform the development of national guidance analogous to NICE guidance, our model does not capture variations between health authorities in provision or efficiency of implementation of the prioritization rules. However, the model can be applied to smaller geographical units that can be of use in decentralized health systems (for example, Italy or Canada). With regards to single hospitals capacity planning, while national guidance could in principle be mandatory, we would argue that—in the case of our modeling—localities should have autonomy to adapt and refine recommendations depending on the opportunities and constraints offered by local service delivery systems and population needs if, for

example, they have excess capacity for certain treatments that would otherwise remain unused.

On the other hand, we do not envisage our model being used at a hospital level, as this would require detailed modeling of hospital-level capacity (by treatment) and patient-level constraints, which in turn would require discrete decision variables that would lead to much harder optimization problems.

This study has several limitations. The impact of COVID-19 on staff shortages and infection-control measures (for example, ward closures) are not modeled, which probably underestimates its impact on hospital capacity. Referrals for elective care are assumed to remain constant at pre-pandemic levels, but it is likely that primary care attendance and referrals were reduced during the pandemic³⁵. This has two implications for the analysis: (1) we do not account for YLL due to reduced care-seeking behavior by patients; (2) the types of referral may differ because of the pandemic, thus impacting costs and YLL of policies.

Patients within each group are considered homogenous in clinical severity and disease progression, and we do not consider competing risks that could change prioritization rules. Also, patients scheduled

Fig. 4 | Comparison of elective and emergency admissions between OS and SP1 under various scenarios. Each dot represents the difference between the number of admissions under SP1 and OS over the 52-week planning horizon for each ICD-10 diagnosis code. This is calculated as $\text{Difference} = \text{Admissions}_{\text{OS}} - \text{Admissions}_{\text{SP1}}$. The vertical black dashed lines represent no difference. Dark purple dots represent elective admissions. Yellow dots represent emergency admissions. Supplementary Fig. 7 presents this information for SP2-4. ICD-10 codes included are A00-B99 (certain infectious and parasitic diseases), C00-D48 (neoplasms), E00-E90 (endocrine, nutritional and metabolic diseases), F00-F99 (mental and behavioral disorders), G00-G99 (diseases of the nervous system), H00-H59 (diseases of the eye and adnexa), I00-I99 (diseases of the circulatory system), J00-J99 (diseases of the respiratory system), K00-K93 (diseases of the digestive system), L00-L99 (diseases of the skin and subcutaneous tissue), M00-M99 (diseases of the musculoskeletal system and connective tissue), N00-N99 (diseases of the genitourinary system), O00-O99 (pregnancy, childbirth and the puerperium), R00-R99 (symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified), S00-T98 (injury, poisoning and certain other consequences of external causes) and V01-Y98 (external causes of morbidity and mortality). The elective bundle is composed of ICD codes A00-B99, E00-E90, F00-F99, H60-H95 (diseases of the ear and mastoid process), O00-O99, P00-P96 and Q00-Q99 (congenital malformations, deformations and chromosomal abnormalities). The emergency bundle is composed of ICD codes D50-D89 (disease of blood, blood-forming organs, and certain disorders involving the immune mechanism), H00-H59, H60-H95, P00-P96, Q00-Q99 and Z00-Z99 (factors influencing health status and contact with health services).

for elective care may be heterogeneous on disease progression and probability of dying. Without an indicator of severity valid across ICDs, we account for latent severity by modeling transition probabilities as a function of waiting times and through some patients being forecasted to need care earlier than others based on historical data. Thus, this analysis may underestimate costs and YLL.

Some patients who have been scheduled for elective care may subsequently die of a hospital-acquired COVID-19 infection; this is not accounted for in our study. If granular data were available,

this could be easily accounted for in the modeling of the transition probabilities. At present, the significance of hospital-acquired infections in the context of our findings is unclear. On the one hand, one could argue that due to capacity shortages, hospitals would use elective wards to admit more emergency patients with COVID-19, thus possibly increasing the risk of hospital-acquired infections. On the other hand, many patients with COVID-19 in the United Kingdom have been allocated to dedicated hospitals (for example, the Nightingale and field hospitals), and the infection-control

Fig. 5 | CC bed utilization by patient group for the OS and SP1 over the 52-week planning horizon. Comparison of bed utilization for the OS (left column) and SP1 (right column) under four scenarios (each row represents one scenario) over the 52-week planning horizon. Each subplot reports, using ICD-10 diagnosis code, the six patient groups with the highest bed utilization rates for both emergency (with red-blue tones) and planned (with gray tones) procedures. Emergency patient groups that cannot be accommodated in the critical care units are reported above the horizontal critical care capacity line. ICD codes included are C00–D48 (neoplasms), I00–I99 (diseases of the circulatory system) and K00–K93 (diseases of the digestive system). Bundle 1 includes ICD codes A00–B99 (certain infectious and parasitic diseases), E00–E90 (endocrine, nutritional and metabolic diseases), F00–F99 (mental and behavioral disorders), H60–H95 (diseases of the ear and mastoid process), O00–O99 (pregnancy, childbirth and the puerperium), P00–P96 (certain conditions originating in the perinatal period) and Q00–Q99 (congenital malformations, deformations and chromosomal abnormalities). Ages 1, 2 and 3 correspond to age brackets <25, 25–64 and 64+, respectively.

programs and protocols developed during the pandemic have helped to reduce hospital COVID-19 transmissions³⁶, thus mitigating the risk of nosocomial COVID-19 in hospitalized patients.

We do not consider treatment preferences of patients, the public or medical professionals³⁷. While our model largely disregards the impact on health inequalities, it embraces the equity principle

that a life year gained is of equal value, regardless of who receives it, which is akin to what is considered by NICE for priority setting. Socioeconomic deprivation and minority ethnicity is associated with increased COVID-19 mortality risk^{38,39}. Further research is needed to ensure that OS do not inadvertently increase health inequalities, and are acceptable to clinicians, patients and the public. If data are available, the model can incorporate health inequalities.

Our objective function minimizes YLL because that is aligned with many prioritization policies routinely used by the NHS in England. For example, the underlying assumption in NICE guidance, followed also internationally, is that the limited NHS budget should be used to maximize health outcomes, albeit using quality adjusted life years (QALYs) rather than YLL. Although QALYs are used by many health systems for economic evaluations of health treatments, data are not routinely available for all diseases and health states for patients in hospital care settings. If data were available, it would in principle be possible to adapt the model to any other objective function, including the maximization of QALYs as well as any other equity considerations.

While these are important caveats that can impact the estimated YLL and costs, they are likely to affect OS and SP in a similar way, thus not impacting their comparison.

The presented model shows that changes in prioritization rules can minimize the detrimental health impact of unprecedented hospital capacity shortages during the pandemic. It operationalizes the principles of best use of limited resources underlying the management of the English health system^{40,41}. More generally, the model is of relevance to health systems globally seeking to prioritize hospital care, substantially improving on short-sighted measures that focus on patients with COVID-19 to the detriment of the health of other patients.

Methods

Optimized schedules. OS are obtained from the solution of an LP model. In the following, we present an abridged model that contains the key decision variables, constraints and input parameters (cf. Supplementary Table 2) of the LP formulation. A more detailed description of the entire optimization model is relegated to Supplementary Section 1.1.

Our LP model can be interpreted as a multiperiod inventory model with five classes of non-negative decision variables: (1) w , tracking patients waiting for care; (2) z , recording which patients are admitted to hospital in a particular week; (3) z' , tracking the initial state of admitted patients (G&A, CC or death because of denied care due to capacity shortages); (4) y , tracking the state of hospitalized patients; (5) x , mapping the transitions of patients across states $s, s' \in \mathcal{S} = \{G, C, G^*, H, D\}$, where G, C and G^* correspond to patients in G&A, CC and G&A after having been denied CC, respectively, while H and D record the recovered and deceased patients.

The abridged LP formulation reads as follows.

$$\min \text{YLL} = \sum_{t \in \mathcal{T}} \sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}} \lambda_p \left(y_{tpa}^D + z_{tpa}^{D'} \right) \quad (1)$$

$$\text{s.t. } w_{t+1,p} = \phi_{tp}^n + \left(1 - \pi_{w,p}^e\right) w_{tp} - z_{tpn} \quad \forall t \neq T, \forall p \quad (2)$$

$$z_{tpe} + z_{tpe}^{D'} = \phi_{tp}^e + \pi_{w,p}^e w_{tp} \quad \forall t \neq T, \forall p \quad (3)$$

$$z_{tpa}^{fs} = z_{tpa} \pi_{z,tpa}^s \quad \forall t \neq T, \forall p, \forall s \in \{G, C\}, \forall a \quad (4)$$

$$y_{t+1,pa}^s = \sum_{s' \in \{H,D\}} x_{tpa}^{s's} \quad \forall t \neq T, \forall p, \forall s, \forall a \quad (5)$$

$$x_{tpa}^{s's} = \pi_{y,pa}^{s's} \left(y_{tpa}^s + z_{tpa}^{s'} \right) \quad \forall t \neq T, \forall p, \forall s' \in \{G, C, G^*\}, \forall s \in \{G, H, D\}, \forall a \quad (6)$$

$$x_{tpa}^{sC} + x_{tpa}^{sG^*} = \pi_{y,pa}^{sC} \left(y_{tpa}^s + z_{tpa}^{s'} \right) \quad \forall t \neq T, \forall p, \forall s \in \{G, C, G^*\}, \forall a \quad (7)$$

The model minimizes the total YLL over a 52-week planning horizon $t \in \mathcal{T} = \{1, \dots, 52\}$ across all patient groups $p \in \mathcal{P}$, each of which comprises

patients of a particular disease type, age group and admission type $a \in \mathcal{A} = \{n, e\}$, where n and e refer to elective and emergency admissions, respectively. The YLL is calculated as the specific YLL per patient group λ_p multiplied by the number of deaths during care (y_{tpa}^D) and due to denial of care ($z_{tpa}^{D'}$) (equation (1)). Other objective functions, such as the minimization of the healthcare costs, or combinations of YLL and healthcare costs, can be readily employed by modifying equation (1) accordingly.

At the beginning of week t , an exogenous inflow of patients in need of care (ϕ_{tp}^a) is observed for every patient group $p \in \mathcal{P}$ and admission type $a \in \mathcal{A}$. Some of the elective patients are admitted to hospital during week t (z_{tpn}), while the other elective patients remain in the waiting queue or require emergency care (with probability $\pi_{w,p}^e$) (equation (2)). Incoming emergency patients (ϕ_{tp}^e) are always admitted if capacity is available. In case of capacity shortages, admission to hospital might be denied to patients in need of emergency care; in the LP model, we conservatively assume that these patients ($z_{tpe}^{D'}$) die due to the lack of the required emergency treatments (equation (3)) (this assumption is relaxed in the sensitivity analysis).

When admitted to hospital, the fraction $\pi_{z,tpa}^G$ ($\pi_{z,tpa}^C$) of patients requiring a G&A (CC) bed is recorded by the decision variables z_{tpa}^G (z_{tpa}^C) (equation (4)). The number of patients of patient group $p \in \mathcal{P}$ and admission type $a \in \mathcal{A}$ in a given state $s \in \mathcal{S}$ at the end of week t , $y_{t+1,pa}^s$, is equal to the sum of the patients who remained in state s during that week (x_{tpa}^{ss}) and the transitions from other states $s' \in \mathcal{S}$ to s during week t ($x_{tpa}^{s's}$) (equation (5)). Transitions across states evolve according to the transition probabilities $\pi_{y,pa}^{s's}$, and they are recorded by the decision variables $x_{tpa}^{s's}$ (equation (6)). Admission to CC is an exception to this: equation (7) enforces that in case of capacity shortages, CC might be denied to a patient in need; the patient then transitions to a designated G&A state (G^*) where s/he evolves according to a new set of transition probabilities until CC capacity becomes available.

The abridged model is a simplified model of the OS and does not account for the possibility to send patients requiring CC to G&A (the designated state G^*) directly upon admission, it assumes that all admitted patients are hospitalized for at least one week, and it disregards the resource constraints of the health system. Those considerations are accounted for by our OS, and they are presented in Supplementary Section 1.1, which also contains more detailed explanations. The solution time of the LP model scales polynomially in the size of the input data, that is, the number of time periods, the number of patient groups as well as the number of resources. Note in particular that the solution time is unaffected by the number of patients, which makes our model particularly suited to policy advice at a national level.

Model inputs. Our model relies on weekly projections of key estimates over the 52-week horizon: patients requiring elective care (Supplementary Section 3.1), non-COVID-19 patients and patients with COVID-19 requiring emergency care (Supplementary Sections 3.1 and 4); probabilities of transitioning to various states once admitted (for example, discharged, to CC, to G&A, died) (Supplementary Section 5.2), and probabilities of patients waiting for elective care needing emergency care (Supplementary Section 5.1). Additional inputs are: forecasts of the number of patients waiting for elective care; hospitalized patients in week 0; costs and YLL of all patients (Supplementary Sections 6.1 and 6.2). Capacity constraints on the supply side are given by the maximum number of G&A/CC beds and staff (senior and junior doctors, and nurses) and recommended staff-to-bed ratios^{42,43}. To reflect historical bed utilization rates, we assume all available capacity can be used.

Patients in need of care are categorized into groups based on primary diagnosis and age resulting in 42 and 45 disease-age groups for non-COVID-19 elective and emergency admissions, respectively, and 3 groups for COVID-19 emergency admissions (Supplementary Section 2). A binary frailty score is calculated for each patient (Supplementary Section 3)⁴⁴.

Non-COVID-19 patients are forecasted using time-series methods assuming historical hospital bed utilization rates (Supplementary Section 3.1). Patients with COVID-19 are projected with a SEIR model using compartments for three age groups and degrees of severity, hospitalizations and deaths. The basic reproduction number, seed time of the epidemic, start of lockdown and reduction in transmission due to non-pharmaceutical interventions are calibrated to hospital occupancy data (Supplementary Section 4)⁴⁵.

The disease progression of patients waiting for elective care may be such that they transition to the cohort requiring emergency care with a probability estimated as a function of waiting time (days) using a Kaplan–Meier estimator (Supplementary Section 5.1). Once patients are admitted for either elective or emergency care to either G&A or CC, they can transition to any of the following states: (1) remain in their current state (G&A or CC); (2) move to CC or G&A; (3) be discharged alive; or (4) die. The transition probabilities are estimated using multinomial logistic regressions, conditional on waiting time for electives (Supplementary Section 5.2).

We compute mean treatment costs for each patient group by averaging across the individual average cost of every non-COVID-19 patient in English hospitals. As the cost of patients with COVID-19 is yet unknown, we determine their HRGs in three London hospitals using the 2015–2019 national cost schedule

(Supplementary Section 6.1). We calculate aggregate YLL by adding the YLL for all deaths within each age group, averaging age specific remaining life expectancies across all ages within the group (Supplementary Section 6)²⁵.

Ethical approval. The study was approved by the NHS Digital Independent Group Advising on the Release of Data (IGARD) committee (application reference: DARS-NIC-276970). No other ethical approval was required.

Data availability

Due to the data sharing agreement, data cannot be made publicly available with this study. Individual-level data used for the analysis of COVID-19 outcomes are not publicly available and only accessible to selected researchers. Data from Hospital Episode Statistics and Electronic Staff Records are available to all researchers upon making an application to NHS Digital. Applications can be made through the NHS's Data Access Request Service (<https://digital.nhs.uk/services/data-access-request-service-dars>). A&E Attendances and Emergency Admissions data are freely available at <https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/>. Bed Availability and Occupancy data are freely available at <https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/>. Critical Care Bed Capacity data are freely available at <https://www.england.nhs.uk/statistics/statistical-work-areas/critical-care-capacity/>. NHS Cost Collection data are freely available at <https://www.england.nhs.uk/national-cost-collection/>. All model input data are documented in the Supplementary Information and made available open source at https://github.com/HFAnalyticsLab/overflow_analysis. Source Data for Figs. 2–5 and for Supplementary Figs. 3–10 are provided with this paper.

Code availability

The source code for the optimization and simulation models used in this study is available open source^{33,34,46}.

Received: 23 February 2021; Accepted: 15 July 2021;
Published online: 13 August 2021

References

- Birnbaum, M. & Morris, L. With coronavirus exploding in Europe, hospitals calculate how long until they hit capacity. *The Washington Post* (2 November 2020); https://www.washingtonpost.com/world/europe/covid-coronavirus-hospitals-capacity/2020/10/31/fe074b56-188a-11eb-8bda-814ca56e138b_story.html
- Emanuel, E. J. et al. Fair allocation of scarce medical resources in the time of COVID-19. *N. Engl. J. Med.* **382**, 2049–2055 (2020).
- Coronavirus: Weston General Hospital halts admissions. *BBC News* (25 May 2020); <https://www.bbc.co.uk/news/uk-england-somerset-52796589>
- Cullen, L. Covid-19: Elective surgery cancelled at Craigavon Hospital. *BBC News NI* (4 November 2020); <https://www.bbc.co.uk/news/uk-northern-ireland-54801326>
- COVID-19 Rapid Guideline: Critical Care in Adults NICE Guideline No. 159 (National Institute for Health and Care Excellence, 12 February 2021); <https://www.ncbi.nlm.nih.gov/books/NBK566886/>
- Important and Urgent—Next Steps on NHS Response to COVID-19 (NHS Improvement and NHS England, 2020); <https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/urgent-next-steps-on-nhs-response-to-covid-19-letter-simon-stevens.pdf>
- Operating Framework for Urgent and Planned Services in Hospital Settings During COVID-19 (NHS England, 2020); <https://covidlawlab.org/wp-content/uploads/2020/06/Operating-framework-for-urgent-and-planned-services-within-hospitals.pdf>
- Raccomandazioni di Etica Clinica per l'Ammissione a Trattamenti Intensivi e per la Loro Sospensione 207-2011 (Società italiana di anestesia analgesia rianimazione e terapia intensiva, 2020); <https://www.quotidianosanita.it/allegati/allegato2675063.pdf>
- Cullen, L. Coronavirus: three of NI's five health trusts cancel planned surgery. *BBC News NI* (12 November 2020); <https://www.bbc.co.uk/news/uk-northern-ireland-54921936>
- Griffin, S. COVID-19: waiting times in England reach record highs. *Br. Med. J.* **370**, m3557 (2020).
- Rosenbaum, L. The untold toll—the pandemic's effects on patients without COVID-19. *N. Engl. J. Med.* **382**, 2368–2371 (2020).
- Campbell, D. NHS hospital waiting lists could hit 10 million in England this year. *The Guardian* (2020). at <https://www.theguardian.com/society/2020/jun/10/nhs-hospital-waiting-lists-could-hit-10-million-in-england-this-year>
- Davey, M. Covid-19 Victoria: hospitals advised to reduce elective surgery as doctors warn of overcrowding. *The Guardian* (9 July 2020); <https://www.theguardian.com/australia-news/2020/jul/09/covid-19-victoria-hospitals-advised-to-reduce-elective-surgery-as-doctors-warn-of-overcrowding>
- Road to recovery. *NHS Confederation* <https://www.nhsconfed.org/news/2020/06/road-to-recovery> (2020).
- Hunter, D. Trying to “Protect the NHS” in the United Kingdom. *N. Engl. J. Med.* **383**, e136 (2020).
- Déry, J. et al. A systematic review of patient prioritization tools in non-emergency healthcare services. *Syst. Rev.* **9**, 227 (2020).
- Valente, R. et al. How to prioritise and optimise access to elective surgery in the COVID-19 pandemic: an interhospital alliance pathway based on the SWALIS 2020 model. A single-cohort feasibility and pilot study. *SSRN* <https://doi.org/10.2139/ssrn.3622349> (2020).
- Cayirli, T. & Veral, E. Outpatient scheduling in health care: a review of literature. *Prod. Oper. Manag.* **12**, 519–549 (2003).
- Ahmadi-Javid, A., Jalali, Z. & Klassen, K. J. Outpatient appointment systems in healthcare: a review of optimization studies. *Eur. J. Oper. Res.* **258**, 3–34 (2017).
- Cardoen, B., Demeulemeester, E. & Beliën, J. Operating room planning and scheduling: a literature review. *Eur. J. Oper. Res.* **201**, 921–932 (2010).
- Gupta, D. & Denton, B. Appointment scheduling in health care: challenges and opportunities. *IIE Trans.* **40**, 800–819 (2008).
- Statement: The Need for National Guidance on Resource Allocation Decisions in the COVID-19 Pandemic (The Nuffield Council on Bioethics, 2021); <https://www.nuffieldbioethics.org/news/statement-the-need-for-national-guidance-on-resource-allocation-decisions-in-the-covid-19-pandemic>
- Hospital Episode Statistics (HES). *NHS Digital* <https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics> (2020).
- Perez-Guzman, P. N. et al. Clinical characteristics and predictors of outcomes of hospitalized patients with coronavirus disease 2019 in a multiethnic London National Health Service Trust: a retrospective cohort study. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciaa1091> (2020).
- Past and Projected Period and Cohort Life Tables, 2018-Based, UK (Office for National Statistics, 2020); <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/pastandprojecteddatafromtheperiodandcohortlifetables/1981to2068>
- NHS England: national cost collection for the NHS. *NHS Digital* <https://www.england.nhs.uk/national-cost-collection/> (2020).
- Statistics: bed availability and occupancy. *NHS Digital* <https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/> (2020).
- Statistics: critical care bed capacity and urgent operations cancelled. *NHS Digital* <https://www.england.nhs.uk/statistics/statistical-work-areas/critical-care-capacity/> (2020).
- Statistics: A&E attendances and emergency admissions. *NHS Digital* <https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/> (2020).
- McCabe, R. et al. Adapting hospital capacity to meet changing demands during the COVID-19 pandemic. *BMC Med.* **18**, 329 (2020).
- Elective Letter (NHS England, 20 August 2020); <https://www.england.nhs.uk/publication/elective-letter/>
- Campbell, D. & Quinn, B. Mass cancellations of NHS operations inevitable this winter, say doctors. *The Guardian* (23 October 2020); <https://www.theguardian.com/society/2020/oct/23/mass-cancellations-of-nhs-operations-inevitable-this-winter-say-doctors>
- jdaeth274 jdaeth274/overflow_analysis: publication release. *Zenodo* <https://doi.org/10.5281/zenodo.5079551> (2021).
- Moret, S. ImperialCollegeLondon/OptimalScheduling4COVID: v1.0. *Zenodo* <https://doi.org/10.5281/ZENODO.5071220> (2021).
- Watt, T., Firth, Z., Fisher, R., Thorlby, R. & Kelly, E. *Use of Primary Care During the COVID-19 Pandemic* (The Health Foundation, 2020); <https://www.health.org.uk/news-and-comment/charts-and-infographics/use-of-primary-care-during-the-covid-19-pandemic>
- Rhee, C. et al. Incidence of nosocomial COVID-19 in patients hospitalized at a large US academic medical center. *JAMA Netw. Open* **3**, e2020498 (2020).
- COVID-19 Documents (The Federation of Surgical Specialty Associations, 2020); https://fssa.org.uk/covid-19_documents.aspx
- Deaths Involving COVID-19 by Local Area and Socioeconomic Deprivation: Deaths Occurring Between 1 March and 31 July 2020 (Office for National Statistics, 2020); <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19bylocalareasanddeprivation/deathsoccurringbetween1marchand31july2020>
- Updating Ethnic Contrasts in Deaths Involving the Coronavirus (COVID-19), England and Wales: Deaths Occurring 2 March to 28 July 2020 (Office for National Statistics, 2020); <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindethsinvolveingthecoronaviruscovid19englandandwales/deathsoccurring2marchto28july2020>
- Next steps on the NHS five year forward view. *Br. Med. J.* <https://doi.org/10.1136/bmj.j1678> (2017).
- Our Principles (NICE, 2020); <https://www.nice.org.uk/about/who-we-are/our-principles>
- Safe Medical Staffing. (Royal College of Physicians, 2020); <https://www.rcplondon.ac.uk/projects/outputs/safe-medical-staffing>

43. *Staffing Levels: Advice Guides* (Royal College of Nursing, 2020); <https://www.rcn.org.uk/get-help/rcn-advice/staffing-levels>
44. Soong, J. et al. Quantifying the prevalence of frailty in English hospitals. *BMJ Open* 5, 8456 (2015).
45. Statistics: COVID-19 hospital activity. *NHS Digital* <https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/> (2020)
46. D'Aeth, J. et al. Optimal scheduling for COVID [source code]. *Code Ocean* <https://doi.org/10.24433/CO.0290795.v1> (2021).

Acknowledgements

We thank the Data Management Team at the Health Foundation for their work to prepare the data extract and manage information governance. This work uses data provided by patients and collected by the NHS as part of their care and support. We also thank C. Donnelly, S. van Esland and S. Riley for comments and suggestions. G.F. acknowledges funding from Jan Wallanders and Tom Hedelius Foundation and the Tore Browaldh Foundation, funding from MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1), jointly funded by the UK Medical Research Council (MRC) and the UK Foreign, Commonwealth and Development Office (FCDO), under the MRC/FCDO Concordat agreement, part of the EDCTP2 program supported by the European Union; and acknowledges funding by Community Jameel. K.H. and N.F. were partially funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Modelling and Health Economics, a partnership between Public Health England, Imperial College London and London School of Hygiene and Tropical Medicine (grant code NIHR200908); and also acknowledge funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1) and by the Community Jameel. P.N.P.-G. acknowledges funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1; and by the Community Jameel. D.H. acknowledges part funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1). J.C.D. acknowledges funding from the Wellcome Trust (reference 102169/Z/13/Z). S.M. acknowledges partial support from the Swiss National Science Foundation (SNSF) under grant number P2ELP2_188028. S.G. was funded by the Imperial College President's PhD Scholarship. F.G. and S.R.D. were funded by the Health Foundation as part of core staff member activity. This research was funded in whole, or in part, by the Wellcome Trust (grant number 102169/Z/13/Z). For the purpose of open access, the author (J.C.D.) has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

All authors contributed to all stages of the work and to the original draft, review and editing. M.M. coordinated the different components of the work. J.C.D., S.G.,

F.G., D.H., E.K., K.L., S.M. and D.R. (listed in alphabetical order) contributed equally as first authors of this work and carried out the analyses. S.G., E.K. and S.M. were responsible for the modeling, optimization and simulation developments. F.G., D.H., D.R., K.L. and P.N.P.-G. were the data curators. M.M., G.F. and W.W. are senior authors and contributed equally. M.M. is the guarantor of the paper and accepts full responsibility for the overall content. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Competing interests

All authors have completed the Unified Competing Interest form (available on request from the corresponding author). K.L. is employed by Flatiron Health but employment started after the research finished. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43588-021-00111-1>.

Correspondence and requests for materials should be addressed to M.M.

Peer review information *Nature Computational Science* thanks the anonymous reviewers for their contribution to the peer review of this work. Handling editor: Ananya Rastogi, in collaboration with the *Nature Computational Science* team.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021