Building a Bayesian decision support system for evaluating COVID-19 countermeasure strategies

Peter Strong\textsuperscript{a}, Aditi Shenvi\textsuperscript{a}, Xuewen Yu\textsuperscript{b}, K. Nadia Papamichael\textsuperscript{c,d}, Henry P. Wynn\textsuperscript{d,e} and Jim Q. Smith\textsuperscript{b,d}\textsuperscript{f}

\textsuperscript{a}Centre for Complexity Science, University of Warwick, Coventry, UK; \textsuperscript{b}Department of Statistics, University of Warwick, Coventry, UK; \textsuperscript{c}Alliance Manchester Business School, The University of Manchester, Manchester, UK; \textsuperscript{d}The Alan Turing Institute, The Alan Turing Institute, London, UK; \textsuperscript{e}London School of Economics and Political Science, London, UK.

1. Introduction

Major disasters and incidents such as pandemics, nuclear disasters, volcano eruptions and tsunamis can impact the short-term and long-term well-being and health of people, nature and economies (see, e.g. Boyd et al., 2014; Jordá et al., 2020; Ohtsuru et al., 2015; Tamura et al., 2000). Decision-making in such crises involves balancing several complex factors, where each factor’s priority may be a subjective judgement (Smith, 2010). Such decision-making is further complicated by the decision centre (DC) – an individual or a group of people, such as a government committee, who must have the authority to enact the decisions made – receiving information from multiple sources. Such information is either factual or narrative in nature, and also may be noisy, uncertain and incomplete (Bakker et al., 2019). Bayesian decision analysis has been shown to enable principled decision-making in such complex, multi-faceted problems (French, 1996; Kurth et al., 2017; Moffat & Witty, 2002; Smith, 2010).

This paper focuses on implementing the well-developed principles of Bayesian multi-criteria decision analysis on the ongoing COVID-19 pandemic which first emerged in Wuhan, China in late 2019 (World Health Organization, 2020). Using the framework presented in this paper, a DC, such as the government, can combine streaming uncertain information they receive about the current and potential effects of the pandemic on the various aspects of society in the form of expected utility scores to decide among a finite set of countermeasure strategies. The study within this paper was undertaken between June and October 2020.

In March 2020, most countries across Europe, including the UK, imposed lockdowns on regional or national levels in an effort to control the spread of the COVID-19 virus, prevent overburdening their healthcare systems and to buy more time to enable researchers to learn more about the disease (Sridhar, 2020). However, lockdowns have been economically damaging, particularly for hard-hit sectors such as aviation, tourism and hospitality (Brien & Harari, 2020). The UK was in recession in the first two quarters of 2020 with early signs of recovery in the third quarter when lockdown restrictions were relaxed (PwC UK, 2020). As of 18th October 2020,
9.6 million jobs have been furloughed under the UK government's job retention scheme (HM Revenue & Customs & UK, 2020). Following the relaxation of initial national lockdown measures in summer of 2020, there have been several localised outbreaks of COVID-19 across the UK, resulting in local lockdowns. The government, keen to avoid another national lockdown to protect the economy, has been under scrutiny regarding the efficacy and frequent changes of the measures introduced under local lockdowns (BBC: Rachel Schraer & Ben Butcher, 2020). In this “second wave”, the government has again faced a challenging balancing act: managing the public health impacts of COVID-19 on one hand, its economic and social impacts on the other. The research presented in this paper was completed against this backdrop; it reached completion in October 2020, when COVID-19 cases were increasing rapidly and the government was on the brink of announcing a second lockdown.

To compare the health, social and economic impacts of the candidate countermeasure strategies, it is necessary to evaluate them on a comparable scale. Typically, health and social impacts are measured in terms of health-adjusted life-years (QALYs) (Miles et al., 2020; Zala et al., 2020) or wellbeing-adjusted life-years (WELLBYs) (De Neve et al., 2020; Layard et al., 2020) as caused directly by the virus itself or indirectly by the countermeasures implemented to tackle it; the economic impacts are measured in terms of difference in realised GDP and empirical or projected GDP, or sometimes as number of jobs in various sectors, consumer spending growth or business investment growth (PwC UK, 2020). Further, these impacts often move in opposite directions in response to any given countermeasure. Hence, a common metric is essential for their comparison (Layard et al., 2020).

Within a Bayesian decision analysis, this common metric is defined through an expected utility score given to each of the health, social and economic attributes. For each attribute, we define a function to convert the actual recorded or estimated measurements (e.g. number of deaths or first order difference in the GDP) into a common metric (here, utilities). Next, by combining these expected utility scores using a specified utility function, we arrive at an expected utility score for each countermeasure strategy. However, the exercise of arriving at an expected utility score for each strategy is complicated by two factors: (1) the specification of the DC’s utility function may involve uncertain estimates of future events (e.g. predicted number of COVID-19 deaths under a complete lockdown), and (2) individuals within a DC may be inclined to prioritise the importance of the various impacts in different ways (e.g. some individuals may view the impact on GDP as more important than others). Consequently, there may be no consensus within a DC as to how efficacious each policy may be.

In this paper, we argue that, in light of the aforementioned challenges, it is essential that any decisions made by policymakers are done after due consideration and analysis within a systematic decision support framework with proper treatment of the associated uncertainties (French, 1995). Such a framework is a defensible and powerful continuous assessment tool and allows policymakers to feed information (for example, the statistics and estimates the UK government obtains from various sources, including its Scientific Advisory Group for Emergencies (SAGE)), and their priorities into the model when it comes to decision making. At the time of writing, the COVID-19 decision support tools in the literature (discussed in Section 2) fall short on providing such a systematic and statistically sound framework that any DC could utilise to compare their options under their self-determined constraints. However, established Bayesian methodologies have been developed over a number of decades to support DCs, capable of respecting both the preferences of the centre and the uncertainties around predictions that inform such decisions, including methods that help compare a variety of different priority weightings. This paper outlines how such a Bayesian decision analysis can help evaluate the efficacy of different COVID-19 countermeasure strategies.

2. Related research

Bayesian decision analyses have been successfully adapted and implemented for various applications, including nuclear disaster support (Geldermann et al., 2009), food security management (Barons et al., 2020), industrial risk management (Rikalovic et al., 2014) and environmental planning (Mattiussi et al., 2014). For more Bayesian and frequentist applications of decision support systems, see Eom and Kim (2006). The COVID-19 threat is fast moving, driven not only by the stochastic spread of the virus but also the dynamic control exerted by a government on the activities of its agents and the general population to the unfolding crisis. Any support tool for managing this crisis needs to acknowledge that, while governments may need to switch between lockdown regimes to increase or decrease the severity of social restrictions in response to the virus prevalence, the general population may tire of the uncertainty in their lives attributable to constantly changing restrictions. In this paper, we draw from previous work (Leonelli & Smith, 2015) to demonstrate how to create a multi-attribute decision
support system (DSS) under a Bayesian approach to address COVID-19 decision making challenges.

We note here that a Bayesian approach has been used for other COVID-19 studies (Dehning et al., 2020; Mbuvha & Marwala, 2020; Neil et al., 2020; Verma et al., 2020), although not for the purpose of providing a decision support framework. Finally, note that in many situations, historical data alone may be inadequate for estimating aggregate utilities from the various attributes. Continuously generated data for our recent and evolving COVID-19 situation, along with empirical data from past epidemics and pandemics are useful to a certain extent but are not sufficient. In this case, it is useful to elicit the required estimates and their associated uncertainties through discussions among a panel of domain experts. The iterative nature of such discussions and the development of the corresponding DSS are detailed in Barons et al. (2018). The structure and estimates of such a DSS would need to be repeatedly reviewed as more information comes to light (e.g. development of a vaccine or new information on immunity from the disease). Typically, the iterative improvements to the DSS are performed until it is deemed to be requisite, i.e. the DC is content that the structure of the DSS is as required (Phillips, 1984). Given the urgency of the issue at hand, it is prudent to begin the process of establishing such a requisite DSS by first developing its framework to address the new challenges of the COVID-19 decision making problem. As more information and data come in, the DSS can be appropriately modified within the framework and the data can be fed into it. Thus our contribution here is the framework for a multi-attribute Bayesian DSS for COVID-19.

There have been several studies focused upon specific aspects of the COVID-19 decision-making problem such as impacts on mortality and poverty (Decerf et al., 2020), a cost-benefit analyses of a lockdown (Layard et al., 2020; Miles et al., 2020), and the impact of specific countermeasure strategies (Karnon, 2020; Lander, 2020; Peto et al., 2020). However, none of these studies adopted a Bayesian approach which we believe is essential in this case as it not only supports uncertainty handling but also gives the government a more transparent and auditable tool. Additionally, these studies look at specific cases but do not describe their general framework. This makes it hard to adapt the existing work as new information comes in. We note here that clinical decision support tools for COVID-19 (e.g. Liu et al., 2020b; McRae et al., 2020; Reeves et al., 2020; Wu et al., 2020) are beyond the scope of discussion for this paper.

In the OR literature, Bayesian approaches have been used to model probabilities (see e.g. Vargo & Cogill, 2015; Zafari & Soyer, 2020). There have been recent calls however, for developing Bayesian subjective utility models to address deep uncertainty in evolving decision contexts (French, 1995; French, 2020). This paper addresses this by developing a framework for providing decision support in an evolving and complex decision context.

3. A COVID-19 decision support framework

In this section, we describe how to construct a framework for a DSS to be used for examining the efficacy of different COVID-19 countermeasures within a Bayesian decision analysis. We assume that the DC satisfies the following key requirements: (1) they must agree on a single agreed rationale for their stated beliefs, and (2) the preferences and any elicited expert judgement used is adopted as their own.

In this paper, we use the definition of a DSS from French et al. (2009) that, “A decision support system is a computer-based system that supports the decision-making process, helping [DCs] to understand the problem before them and to form and explore the implications of their judgements, and hence to make a decision based upon understanding”.

It is convenient to break the process of creating a DSS down into four phases:

3.1. Elicitation of the class of strategies we might consider and the attributes of the utility function

We start by considering the types of strategy whose effectiveness we need to assess. In this framework, for the purpose of this paper, a strategy is defined by the regimes that might be imposed (such as a lockdown) and the thresholds (based on the system, which when reached leads to a switching of regimes). These strategies are designed to control the behaviour of the population and are typically tiered by the level of stringency these represent, i.e. the amount of disruption they cause to the normal life of the population.

The attributes of the utility function are the different features of interest that each strategy affects. We define \( A = \{A_1, A_2, ..., A_m\} \) as the set of attributes of interest, with \( a_i \) the value that attributes \( A_i \).

In the context of decision support for COVID-19, attributes which may be of interest include: the years of life saved across the population from avoiding infection by the virus; the years saved through timely medical examinations, for example through the detection of cancer and the encouragement to report to Accident and Emergency departments (A&E) when exhibiting symptoms of a stroke or heart attack. A further threat to survival will be the response to increased poverty as categorised by the
distribution of life expectancy given movements in the social class across the population induced by, for example, less effective education, unemployment or reduced employment activities. Note that all these attributes can be measured in units of years of life saved. On top of these, we may need to consider further attributes that define the quality of life of the population or its wellbeing, such as measures of depression and anxiety or economic hardship.

3.2. Elicitation of a quantification of the DC’s marginal utilities of these attributes and the criterion weights

A subjective expected utility analysis consists of two components: a utility function \( U(a) \) over a set of measured attributes \( a = (a_1, a_2, \ldots, a_m) \) and a set of multivariate probability densities \( \{ p_s(a) : s \in S \} \) into the future associated with each possible countermeasure strategy. Whilst the latter would be provided by the appropriate domain experts working with statisticians and mathematical modellers, the utility function \( U(a) \) needs to be elicited from the DC to reflect how they intend to frame their objectives and prioritise risk. More precisely, if \( a^- \) and \( a^+ \) denote what the DC perceive to be respectively the worst and best credible outcomes then for each vector of outcomes \( a, \{ a : a^- \leq a \leq a^+ \} \), \( U(a) \) is an increasing linear function of the probability \( q(a) \) where the DC finds the outcome \( a \) with certainty equally preferable to a hypothetical situation where they are faced with obtaining the best possible outcome \( a^+ \) with probability \( q(a) \) and the worst \( a^- \) with probability \( 1 - q(a) \). In this way, aspects of the DC’s risk aversion can be captured (see Papamichail & French, 2003; Smith, 2010).

On the basis of certain basic axioms (see Smith, 2010), various ways have been devised to indirectly elicit these preferences efficiently, effectively and specifically with less biases. One assumption that is often made, and can be checked against a DC’s expressed preferences, is that the DC has value independent attributes. The attributes \( a \) are said to be value independent if the DC always finds two strategies \( s_1 \) and \( s_2 \) (leading to densities over outcomes \( \pi_1(a) \) and \( \pi_2(a) \) respectively) equally preferable whenever the marginal distributions over \( \pi_1(a) \) and \( \pi_2(a) \) for each component attribute in \( a \) are the same. It can then be shown that we can find positive criterion weights \( (k_1, k_2, \ldots, k_m) \), \( \sum_{i=1}^{m} k_i = 1 \) such that \( U(a) \) can be written

\[
U(a) = \sum_{i=1}^{m} k_i U_i(a_i)
\]

where \( U_i(a_i) \) – called the marginal utility of \( a_i \) – is an increasing function of its argument, and \( k_i \geq 0, i = 1, 2, \ldots, m \). Whether the attributes satisfy value independence is not an inherent feature of the attributes but is a subjective judgement of the DC. It has been found that in practice, provided the attribute vector \( a \) is carefully defined, this simple form well approximates the DC’s actual utility function in the vast majority of analyses. The advantage of making this assumption is that \( U(a) \) is then much easier to elicit and the output of the analysis much more transparent and easy to explain: see (Smith, 2010) for a long discussion of these points. In particular, the criterion weights of the different component attributes can be chosen to reflect their relative importance. On the other hand, the form of the marginal utilities can be chosen to reflect the extent to which a DC considers an outcome good or bad relative to its extremes for each single attribute in turn. Henceforth, in this paper, we will assume the DC has value independent attributes.

3.3. Building a probabilistic model of each attribute and performing a Bayesian analysis which combines available data with probabilistic expert judgements

This is the process of obtaining \( p(a_i|s) \), the probability of attribute \( i \) having value \( a_i \), given strategy \( s \). In the context of COVID-19, this involves obtaining a probabilistic model for the number of deaths based on the strategy taken. This requires some sophistication since the spread of the virus at any time is complex and uncertain as it is a function of both the regime that a strategy is in at that time and the latent state of the disease. Therefore, any model informing the DC needs to be able to accommodate expert judgements, epidemiological modelling and data, all synthesised through Bayesian techniques; where in complex cases these techniques would tend to be computational.

3.4. Calculating a final score on each attribute as an expectation of its utility function

These scores are weighted by the elicited criterion weights to provide expected utility scores for each of the strategies considered. Note that there is always an explanatory rationale to whether or not a strategy scores well based on each of the different individual attribute scores and the predictions of the uncertain consequences of each strategy.

As the attributes chosen are value independent (Insua & French, 2010; Keeney & Raiffa, 1993), the subjective expected utility score \( \bar{U}(s) \) of each strategy \( s \) over the \( m \) attributes \( a_1, a_2, \ldots, a_m \) with respective marginal utilities \( U_1, U_2, \ldots, U_m \) and the associated criterion weights \( k_1, k_2, \ldots, k_m \) elicited in phase 2, is given by
Here, the marginal expected utility scores for each attribute \(j\) given strategy \(s\) are given by

\[
\bar{U}(s) = \sum_{i=1}^{m} k_i \bar{U}_i(s).
\]

4. A simplified Bayesian analysis of COVID-19 strategies

In this section, we give a simple illustration of how a formal Bayesian multi-criteria decision analysis, as described in Section 3, could have been used to help DCs weigh the efficacy of different options open to them. Typically, marginal utility functions would need to be elicited. We have mentioned above that these would typically be risk-averse, which would have the effect of reducing expected utility scores when the outcomes of associated strategies are more uncertain. However, because our main focus in this example is how the Bayesian multi-criteria decision analysis balances the efficacy achieved associated with a given attribute is always uncertain throughout the decision making process.

4.1. Class of strategies and attributes

4.1.1. Regimes

In practice, there is a large set of different regimes to consider, including potentially all possible combinations of preventative measures such as: closing schools, making mask wearing compulsory and promoting working from home. Here, we have chosen to limit our example to 3 regimes that the strategies will be able to switch between for simplicity. These loosely correspond to a regime implemented so far by the UK government. The three regimes are:

- \(r_0\) – No lockdown: No restrictions.
- \(r_1\) – Partial lockdown: Return to work and school, non-essential businesses open with 1m + distancing.
- \(r_2\) – Complete lockdown: Work from home in effect, ban on non-essential interactions, complete closure of schools and non-essential businesses.

The first two regimes are analogous to the UK policies in effect on 26th March and 4th July respectively. A strategy determines when and how the transitions between these regimes occur, with the possibility of no transitions allowed.

4.1.2. Strategies

As the DC will be responding to the spread of the disease in a population which is susceptible to shocks such as large outbreaks, strategies considered should naturally include those which include switching between these three regimes.

In this example, we will consider several strategies that have different thresholds that give the transitions between regimes. At the time of writing, the thresholds on what measures the government was using to inform their strategies were not publicly available. In lieu of this, the transition thresholds for our example were based around the decisions made by England if they were based on the metrics defined in our example.

In practice, a DC would likely be interested in a much wider range of strategies. However, in our example, for ease of analysis, we will narrow down the set of strategies that we are interested in by putting the following conditions on them:

- At most, one switch between regimes can take place in a week.
- Within the timescale of the simulation, once restrictions have began, we do not return to no restrictions.
- The transitions between regimes here are set such that they are dependent on the cumulative reported deaths, which is assumed to be known and the reported proportion of the population infected, which is assumed to be known with a delay of a week. In practice, this could be estimated from the proportion of positive tests from a sample of the population with the information being made available with a delay.

Thus, the state-space diagram for each countermeasure strategy considered corresponds to one of the figures in Figure 1.
Below, we outline the different conditions that a strategy might have for switching between regimes.

**Initial lockdown** ($r_0 \rightarrow r_1/r_0 \rightarrow r_2$): Given that we are in regime $r_0$, a transition to $r_1$ or $r_2$ occurs when the cumulative deaths are larger than some chosen critical value $L$. We consider three choices of $L$:

- Earlier lockdown: $L_1 = 100$;
- Later lockdown: $L_3 = 500$.

Note that since the population can only transition out of $r_0$ at most once, there is no need to set out different conditions for transitioning to $r_1$ and $r_2$. Instead, each strategy specifies whether its transition out of $r_0$ is to $r_1$ or $r_2$.

**Tightening of lockdown** ($r_1 \rightarrow r_2$): When a population is under regime $r_1$, a 5% rise in the observed number of cases in a week will cause a transition into $r_2$. As no tightening of lockdown had occurred at the time of this study, the choice of a 5% rise was an arbitrary one for illustrative purposes and, in practice, a range of values would be considered.

**Easing of lockdown** ($r_2 \rightarrow r_1$): Given that a population is currently under regime $r_2$, a transition into $r_1$ occurs when the proportion of infected individuals falls below a certain critical proportion $E$ of the peak proportion of infected individuals since $r_2$ has been in effect. We consider the following choices of $E$:

- No easing: $E_0 = 0$;
- Easing in line with when the UK decided to ease lockdown: $E_1 = 0.12$;
- Quicker easing: $E_2 = 0.3$;
- Very quick easing: $E_3 = 0.5$.

Based on these transition rules, we assume that the initial transition out of $r_0$ is to $r_2$ for all but three of the strategies we consider. A strategy denoted as $L_{E_i}$ has initial transition from $r_0$ to $r_2$ under critical value $L_i$ and allows for easing from $r_2$ to $r_1$ under critical proportion $E_j$ where $i \in \{1,2,3\}$ and $j \in \{0,1,2,3\}$. Further, analogous to “no complete lockdown” strategies, we define three strategies denoted by $L_{E_i}^*$ where the initial transition is from $r_0$ to $r_1$ under critical value $L_i$, $i \in \{1,2,3\}$. Under these three strategies, once regime $r_1$ is effected, it remains in place until the end of the simulations with no transition into $r_2$ permitted.

### 4.1.3 Attributes

We will consider the attributes shown in Figure 2 that shows the attribute tree for our example. Since we are only considering health-related attributes, all the attributes can be measured in terms of aggregate expected life-years lost by the population. In this example, we are interested in the attributes of:

- $a_1$: Life-years lost due to COVID-19.
- $a_2$: Excess life-years lost due to poverty.
- $a_3$: Excess life-years lost due to delayed cancer diagnosis.

This attribute tree can be easily adapted to add other value independent attributes (see Section 3) of interest for the DC such as cost of strategy implementation, economic impacts and social impacts such as quality of life.

### 4.2 Marginal utilities and criterion weights

#### 4.2.1 Marginal utilities

As in this example each of our attributes is in terms of life-years lost, we will take our marginal utility function to be the negative identity function, $U_i(a_i) = -a_i$ \( \forall i \).

The values of $a_i$ are, of course, uncertain. However, because of the assumption of preference independence and this linear utility function for the evaluation of the score, we need only elicit the subjective expectation of this quantity under each strategy. Here, again for simplicity, we have chosen to identify this expectation with the output a stochastic model of the process. We assume that the centre adopts as their expectation those delivered by relevant experts. For a theory justifying when this is an appropriate protocol for a DC to adopt, see Leonelli and Smith (2015) and Smith et al. (2015).

#### 4.2.2 Criterion weights

The criterion weights for these attributes reflect how the DC prioritises them. We compare the effects of
different setting of the attribute weights $k = (k_1, k_2, k_3)$ on the aggregate weighted life-years lost where weight $k_i$ is for attribute $a_i$, $i = 1, 2, 3$. For this example, we will consider the effect of several choices of criterion weights:

- $(1, 0, 0)$: The DC only cares about life-years lost due to COVID-19.
- $(0.5, 0.5, 0)$: The DC only cares about life-years lost due to COVID-19 and delayed cancer diagnosis and weights them equally.
- $(1/3, 1/3, 1/3)$: The DC cares about all 3 causes of life-years lost and cares about them all equally.
- $(0.45, 0.45, 0.1)$: The DC cares about all 3 causes of life-years lost and but life-years lost caused by short and medium term attributes are more important.

4.3. Probabilistic model of each attribute

For each attribute, we will give a probabilistic model giving us the probability of an attribute obtaining each value for each strategy, $p(a_i|s)$.

4.3.1. Covid-19 deaths ($a_1$)

To estimate the deaths from COVID-19 and the proportion of infected individuals in the population, we use a simple discrete-time SIRD compartmental epidemiological model. The states in this model are “Susceptible” (S), “Infectious” (I), “Recovered” (R) and “Dead” (D). The dynamics of this model can be seen in Figure 3 and are described in the supplementary material S1. Clearly, D is an absorbing state and we assume that those that have recovered from COVID-19 are immune to it. We stratify our population in this model by region and age.

In order to capture the uncertain nature of the future, we have modelled uncertainty around the rate of infection with detail given in supplementary material S1.2. By running our stochastic simulation 1000 times we obtained expected values for both the attribute of number of deaths due to COVID-19 and the number of weeks spent in each regime. One of these simulations for each strategy is shown in Figure 4. Note here that in practice the expected numbers used are those provided by epidemiologists using more sophisticated models.

4.3.2. Delayed cancer diagnoses ($a_2$)

In this example, deaths due to delayed diagnoses of cancer are used here as a proxy for deaths that would not have normally occurred, as a composite of missed treatments and patients who have not presented in hospital. This was chosen as a proxy as information was easily available. In practice, this could be improved on by either creating attributes for other excess deaths due to missed treatments and eliciting models of each attribute or, when that is not feasible, updating the utility function for this attribute. The latter could be done simply by assuming that deaths due to delayed cancer diagnosis will form a certain proportion of all of the deaths due to missed treatment. Estimates like this are often necessary in the early stages of a decision support system for an emergency response when there is sparse information about the underlying science.

Sud et al. (2020) evaluate how delays to cancer diagnoses due to the COVID-19 outbreak impact survival and life-years lost. This study gives a nearly linear relationship for delay in cancer diagnoses to extra deaths per age group. Using this relationship, we are able to produce plausible values for the expected total number of deaths arising from this delay. These can be used as benchmark values the DC might use for its subjective expectations of this attribute.

As at the time of this report there is little research done on the impact of partial lockdown on cancer referrals, we will assume that the number of deaths in a partial lockdown is half of those in a total lockdown. In practice, the DC would elicit expert judgement about this value where possible and perform a sensitivity analysis.

4.3.3. Poverty ($a_3$)

Here, we notice that the centre has other experts available who can assess the impact on years of life lost on poverty induced by a strategy.

For example, Decerf et al. (2020) evaluate the impact of the COVID-19 outbreak in terms of number of poverty-years. Poverty-years are calculated using...
COVID-19’s predicted effect on GDP and estimates the number of additional people in poverty as a result (in the UK, 4.37 million). Decerf et al. (2020) conservatively assumes that these individuals will remain poor only for a single year. We use Decerf et al. (2020)’s calculation that 8.8 poverty-years equate to 1 life-year to inform our DSS, which gives us an estimate for the impact of lockdown on life-years lost due to COVID-19.
Further information is given in the supplementary material section S2.

4.4. Final scores calculated

We can now combine this all together to get Equation 3 and assess which strategy maximises the utility score given the criterion weights.

\[
\bar{U}(s) = \sum_{i=1}^{3} k_i \bar{U}_i(s) = \sum_{i=1}^{3} k_i \left( -a_i(s) \right) \tag{3}
\]

Here \( \bar{a}_i(s) \) is the expected value of attribute \( i \) under strategy \( s \). In this example, this is the magnitude of the utility function, \( \bar{U}_i(s) \).

4.5. Results

From Figure 5, we can see that under \( k = (1, 0, 0) \) where only life-years lost due to COVID-19 are of interest, strategies that do not involve a lockdown perform orders of magnitudes worse. Looking at strategies that do involve lockdown, “earlier lockdown” and “slower easing of lockdown” lead to fewer life-years lost. The second scenario under \( k = (0.5, 0.5, 0) \), in which life-years lost due to delayed cancer diagnoses are also considered, we can begin to see the drawbacks of the lockdown with reduced disparity between strategies that include lockdown and those that do not. However, the lockdown strategies remain more desirable. This example also shows little difference between the lockdown strategies that only differed in their easing of lockdown with more significant differences depending on when the lockdown conditions were implemented.

Under \( k = (1/3, 1/3, 1/3) \), which includes life-years lost due to poverty with equal weighting for all attributes, the non-lockdown strategies perform the best. It is interesting to see how small the weight on our long term impacts on health might need to
be before we impose the most stringent of lockdown strategies. For \( k = (0.45, 0.45, 0.1) \), we treat short and medium term years lost as more important than long term years lost. While lockdown strategies generally still perform better, the disparity among all strategies is much less.

We further explore these trade-offs in Figure 6, which shows a Pareto front plotting life-years lost due to poverty against those lost from COVID-19 and cancer combined, assuming equal weighting. We can see how the trade-off in weights leads to different strategies being the most effective, with the weights giving the gradient of the line. This plot can help to reduce the number of strategies considered, as in this example strategies with no easing of lockdown are strictly dominated by strategies with quick easing of lockdown. Thus, there is no weighting of attributes in which we would rather have no easing. We can also find the value \( c \) in \( k^* = (c, c, 1 - 2c) \) for which the DC would swap preferences between strategies that involve lockdown and those that don’t. In this example, that point is at \( c = 0.419 \), where short and medium term attributes are weighted as 2.58 times more important than long term attributes. At \( c \geq 0.419 \), the optimal weighting would be strategies that involve lockdown. For comparing trade-offs between more attributes at the same time, an n-dimensional Pareto front with each axis on its own label should be used. Further analyses of the trade-offs between the attributes using Pareto fronts is given in the supplementary materials S3-S4.

5. Discussion

In this paper, we presented how a DC can create a DSS under which countermeasure strategies for COVID-19 could be evaluated in real-time, using the best expert judgements associated with various strategies that it might be feasible for the DC to enact. This DSS enables the DC to input the attributes they consider to be vital, with attribute weightings reflecting their priorities, to evaluate and choose an appropriate metric by which these attributes can be measured.

The novelty of the framework lies in demonstrating how to provide multi-attribute Bayesian decision support in evolving decision contexts. We have shown that Bayesian decision analysis is an appropriate and holistic approach for addressing uncertainty in dynamic environments as we move from short-term to longer-term considerations. Indeed, the results of a Bayesian subjective utility approach are informative and help decision makers devise action plans (French, 2020).

In the results from our example, we observed the effect of different criterion weights demonstrating their effect on the optimal decision. For illustrative purposes, the analysis presented here was simplified. In order to be operationalised, it would need to be refined with additional attributes as well as a utility function and attribute weightings that reflect the priorities of the DC. Firstly, more attributes – such as hospital admissions, well-being, economic viability and public acceptability – could be considered.

Further, such a DSS could be extended by considering the impact of other countries’ management of COVID-19, caused by a spillover effect. For example, the UK’s economy is likely to be affected by COVID-19 spreading in other countries regardless of the UK’s state of lockdown. Our model implicitly assumes that such further causes of deprivation are additive, whereas the relationship between impacts to domestic employment and international economic variables on the differential scores associated with different options is likely to be more subtle. However, we note that, by embarking on this decision analysis, we are drawn into trying to quantify such issues and to fold these important considerations into the analysis; this DSS gives a framework for introducing the results of studies as these become available.
Another critical issue that could be incorporated would be the capacity of the health provider, e.g. the National Health Service (NHS) in the UK, both with respect to the number of available beds and the number of doctors and nurses available to treat COVID-19 as well as non-COVID-19 patients. Hospitals functioning close to full capacity may lead to increase in the mortality rate of hospitalised COVID-19 patients (Wilde et al., 2021) as well as temporary closure of all non-urgent hospital procedures (COVIDSurg Collaborative, 2020). Thus, if the NHS were to be overwhelmed, the underlying model would need to be adapted to reflect the increased life years lost due to the added pressure. Indeed, in the UK, COVID-19 policy decisions have been influenced by the capacity of the NHS (Guardian: Rowena Mason & Nicola Davis, 2021), and measures such as introducing additional capacity through construction of temporary hospital facilities (known in the UK as Nightingale Hospitals) (NHS, 2020) were effected to prevent overwhelming the NHS. These capacities could be incorporated (1) in deciding the thresholds for changes in regimes and (2) by choosing hospital admissions as an attribute. The utility function of this attribute could be an indicator function that indicates when the hospital capacity is exceeded. Assigning a very large criterion weight to such an attribute would have the effect of making any strategy that would lead to the threshold being passed strictly dominated by other strategies and therefore, not considered.

We also assumed that all the marginal utilities were linear. In contexts like the one above, the DC tends to be risk-averse. The mathematics means that the scores we assign to different options which lead to more uncertain outcomes are penalised. There are two elements that come into play here. The first issue is that the impacts of previously untried strategies are likely to be more uncertain. A risk-averse marginal utility would therefore tend to down-weight the scores of less well-tried options. The second is that attributes which are intrinsically more uncertain will tend to be given less weight across all strategies. For example, the implicit assumption of giving negligible weight to non-poverty related economic effects could be justified if the differential economic effects of one strategy against another were extremely uncertain. However, once the marginal utilities have been elicited, the framework is able to score the different options in light of this and the type of analysis above can be adapted. Further, compliance of the population to flip-flopping between regimes could be accounted for by considering it within the probabilistic predictions of the attributes.

Finally, we have used a naïve predictive model of the progress of the disease. This dynamic model could obviously be improved; see Keeling et al. (2021); Aguilar et al. (2020). In particular, we could choose a stochastic disease model such as Abrams et al. (2020). Assumptions on immunity after recovery might also need to be revised as relevant studies become available; for example Long et al. (2020); Liu et al. (2020a).

In conclusion, further research is needed to extend on this work in all the ways mentioned above. However, we have shown, using as simple example, how a DSS could be created and analysed to assess different countermeasure strategies for COVID-19. We have also shown how this could be done in a real time setting – where there is only basic scientific understanding of such a disease and the impacts of its countermeasures, as witnessed for COVID-19 – where DCs may need to rely on only coarse descriptions given by the experts when many judgements are unknown or unsubstantiated by evidence.

Note
1. The code used to generate this example is available here: https://github.com/peterrhysstrong/COVID-19-DSS

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
Peter Strong was funded by the EPSRC and MRC Centre for Doctoral Training in Mathematics for Real-World Systems (EP/L015374/1). Aditi Shenvi was funded by the University of Warwick Chancellor’s International Scholarship; K.Nadia Papamichail’s and Jim Q. Smith’s work was supported by a Fellowship from the Alan Turing Institute.

ORCID
Peter Strong http://orcid.org/0000-0002-1514-9698
Aditi Shenvi http://orcid.org/0000-0002-2453-9188
Xuewen Yu http://orcid.org/0000-0001-9759-0423
K. Nadia Papamichail http://orcid.org/0000-0001-8307-9620
Henry P. Wynn http://orcid.org/0000-0002-6448-1080
Jim Q. Smith http://orcid.org/0000-0002-9224-5317

References
Abrams, S., Wambua, J., Santermans, E., Willem, L., Kuylen, E., Coletti, P., Libin, P., Faes, C., Petrof, O., & Herzog, S. A. (2020). Modeling the early phase of the Belgian COVID-19 epidemic using a stochastic compartmental model and studying its implied future trajectories. *Epidemics*, 35. https://doi.org/10.1016/j.epidem.2021.100449
McRae, M. P., Simmons, G. W., Christodoulides, N. J., Lu, Z., Kang, S. K., Fenyo, D., Alcorn, T., Dapkins, I. P., Sharif, L., Vurmaz, D., Modak, S. S., Srinivasan, K., Warhadpande, S., Shrivastav, R., & McDevitt, J. T. (2020). Clinical decision support tool and rapid point-of-care platform for determining disease severity in patients with COVID-19. *Lab on a Chip*, 20(12), 2075–2085. https://doi.org/10.1039/D0LC00373E

Miles, D. K., Stedman, M., & Heald, A. H. (2020). Stay at Home, Protect the National Health Service, Save Lives*: A cost benefit analysis of the lockdown in the United Kingdom. *International Journal of Clinical Practice*, 75(3), e13674.

Moffat, J., & Witty, S. (2002). Bayesian decision making and military command and control. *Journal of the Operational Research Society*, 53(7), 709–718. https://doi.org/10.1057/palgrave.jors.2601347

Neil, M., Fenton, N., Osman, M., & McLachlan, S. (2020). Bayesian Network Analysis of COVID-19 data reveals higher Infection Prevalence Rates and lower Fatality Rates than widely reported. *Journal of Risk Research*, 23(7–8), 866–879.

NHS. (2020). NHS to build more nightingale hospitals, as London set for opening.

Ohtsuru, A., Tanigawa, K., Kumagai, A., Niwa, O., Takamura, N., Midorikawa, S., Nollet, K., Yamashita, S., Ohno, H., Chhem, R. K., & Clarke, M. (2015). Nuclear disasters and health: Lessons learned, challenges, and proposals. *The Lancet*, 386(9992), 489–497. https://doi.org/10.1016/S0140-6736(15)60994-1

Papamichail, K. N., & French, S. (2003). Explaining and justifying the advice of a decision support system: A natural language generation approach. *Expert Systems with Applications*, 24(1), 35–48. https://doi.org/10.1016/S0957-4174(02)00081-7

Peto, J., Alwan, N., A., Godfrey, K. M., Burgess, R. A., Hunter, D. J., Riboli, E., Romer, P., Buchan, I., Colbourn, T., Costelloe, C., Davey Smith, G., Elliott, P., Ezzati, M., Gilbert, R., Gilthorpe, M. S., Foy, R., Houstoun, R., Inskip, H., Lawlor, D. A., … Yao, G. L. (2020). Universal weekly testing as the UK COVID-19 lockdown exit strategy. *The Lancet*, 395(10234), 1420–1421. https://doi.org/10.1016/S0140-6736(20)30936-3

Phillips, L. D. (1984). A theory of requisite decision models. *Acta Psychologica*, 56(1-3), 29–48. https://doi.org/10.1016/0001-6918(84)90005-2

PwC UK. (2020). UK Economic update.

Reeves, J. J., Hollandsworth, H. M., Torriani, F. J., Taplitz, R., Abeles, S., Tai-Seale, M., Millen, M., Clay, B. J., & Longhurst, C. A. (2020). Rapid response to COVID-19: Health informatics support for outbreak management in an academic health system. *Journal of the American Medical Informatics Association: JAMIA*, 27(6), 853–859. https://doi.org/10.1093/jamia/ocaa037

Rikalovic, A., Cosic, I., & Lazarevic, D. (2014). GIS based multi-criteria analysis for industrial site selection. *Procedia Engineering*, 69(12), 1054–1063. https://doi.org/10.1016/j.proeng.2014.03.090

Smith, J. Q. (2010). Bayesian decision analysis: Principles and practice. Cambridge University Press.