Simpson's paradox in Covid-19 case fatality rates: a mediation analysis of age-related causal effects

Julius von Kügelgen[*1,2] **Luigi Gresele**[*1,3] **Bernhard Schölkopf** [1]

[1]Max Planck Institute for Intelligent Systems, Tübingen, Germany
[2]Department of Engineering, University of Cambridge, United Kingdom
[3]Max Planck Institute for Biological Cybernetics, Tübingen, Germany
{jvk, luigi.gresele, bs}@tuebingen.mpg.de

Abstract

We point out an instantiation of Simpson's paradox in Covid-19 case fatality rates (CFRs): comparing data of 44,672 cases from China with early reports from Italy (9th March), we find that CFRs are lower in Italy for every age group, but higher overall. This phenomenon is explained by a stark difference in case demographic between the two countries. Using this as a motivating example, we introduce basic concepts from mediation analysis and show how these can be used to quantify different direct and indirect effects when assuming a coarse-grained causal graph involving country, age, and mortality. As a case study, we then investigate total, direct, and indirect (age-mediated) causal effects between different countries and at different points in time. This allows us to separate age-related effects from others unrelated to age, and thus facilitates a more transparent comparison of CFRs across countries throughout the evolution of the Covid-19 pandemic.

1 Introduction

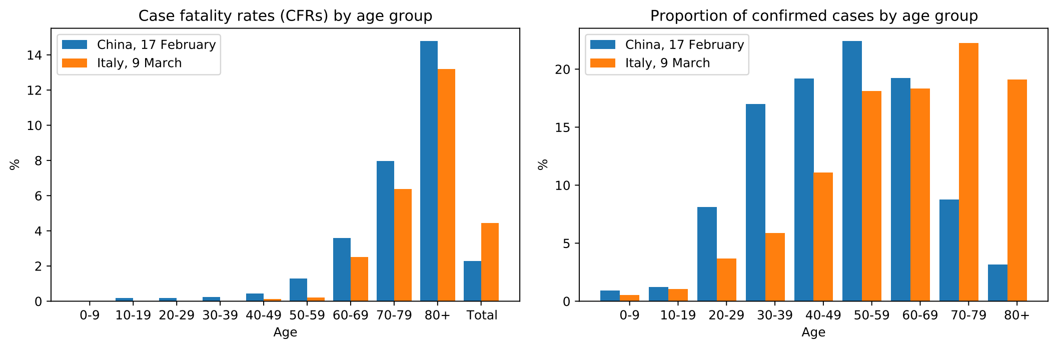
The 2019–20 coronavirus pandemic originates from the SARS-CoV-2 virus, which causes the associated infectious respiratory disease Covid-19 [4]. After an outbreak was identified in Wuhan, China, in December 2019, cases started being reported across multiple countries all over the world, ultimately leading to the World Health Organization (WHO) declaring it a pandemic on 11 March 2020 [26]. As of 3 June 2020, the pandemic led to more than 6.44 million confirmed cases and over 382,000 fatalities across 188 countries [25]. One of the most cited indicators regarding Covid-19 is the reported case fatality rate (CFR), which indicates the proportion of confirmed cases which end fatally. In addition to the *total* CFR, CFRs are often also reported separately *by age* since CFRs differ significantly across different age groups, with older people statistically at higher risk.

In this work, we show how tools from causal inference and, in particular, mediation analysis can be used to interpret case and fatality data related to the Covid-19 pandemic. We motivate our investigation by pointing out what could be a textbook example of Simpson's paradox in comparing CFRs between China and Italy, suggesting opposite conclusions depending on whether the data is analysed in aggregate or age-stratified form (§2). This example illustrates how a traditional statistical analysis provides insufficient understanding of the data, and thus needs to be augmented by additional assumptions about the underlying causal relationships. In §3, we therefore postulate a coarse-grained causal model for comparing age-specific Covid-19 CFR data across different countries. We then review different types of (direct and indirect) causal effects, and motivate them in the context of our assumed model as different questions about Covid-19 fatality in §4.

As one of our contributions, we curated a dataset involving 756,004 confirmed Covid-19 cases and 68,508 fatalities, separated into age groups of 10-year intervals (0–9, 10–19, etc.), reported from 11 different countries from Africa, Asia, Europe and South America and the Diamond Princess cruise ship, which, together with an interactive notebook containing all our analyses, is publicly available

[*]Equal contribution

Preprint. Under review.



(a) (b)

Figure 1: (a) Snapshot of Covid-19 case fatality rates (CFRs) in Italy and China by age group and in aggregated form ("Total"), i.e., including all confirmed cases and fatalities up to the time of reporting (see legend). (b) Proportion of cases included in (a) within each age group. Sources: [27, 10].

at: `github.com/Juliusvk/Covid19-age-related-causal-effects`. We use this dataset, in combination with the proposed coarse-grained model, to perform a case study (§5). Tracing the evolution of direct and indirect effects of country (China or Italy) on mortality, from early March to late May 2020, allows to discover trends—such as a reversal in the sign of the direct effect in mid March that temporally aligns with a reported "collapse" of the health-care system in parts of Italy [1]—which may otherwise remain hidden in the data. Moreover, we compute direct and indirect effects for 132 pairs of countries and thus identify countries whose total CFRs are particularly adversely affected by their case demographic. We further find that the size of indirect effects is strongly correlated with a country's median age, but only weakly with that of direct effects.

While our model is rather simple—and we do not claim to be the first to introduce it—it constitutes, to the best of our knowledge, the first application of causal analysis to better understand the role of mediators such as age in Covid-19 fatality data. We discuss limitations such as issues of selection bias and propose future directions such as the inclusion of additional mediators in §6, and hope that our work may serve as a stepping stone for further studies to gain better insight into the mechanisms underlying the Covid-19 pandemic within a principled and transparent causal framework.

2 Simpson's paradox in comparing CFRs between China and Italy

When comparing Covid-19 CFRs for different age groups (i.e., the proportion of confirmed Covid-19 cases within a given age group which end fatal) reported by the Chinese Center for Disease Control and Prevention [27] with preliminary CFRs from Italy as reported on March 9 by the Italian National Institute of Health [10] a seemingly strange pattern can be observed: *for all age groups,* CFRs *in Italy are lower than those in China, but the total* CFR *in Italy is higher than that in China.* This is illustrated in Fig. 1a (see Table 1 in Appendix A for exact numbers). It constitutes a textbook example of a statistical phenomenon known as *Simpson's paradox* (or *reversal*) which refers to the observation that aggregating data across subpopulations (here, age groups) may yield opposite trends (and thus lead to reversed conclusions) from considering subpopulations separately [24].

How can such a pattern be explained? The key to understanding the phenomenon lies in the fact that we are dealing with *relative* frequencies: the CFRs shown in percent in Fig. 1a are ratios and correspond to the conditional probabilities of fatality given a case from a particular age group and country. However, such percentages conceal the absolute numbers of cases within each age group. Considering these absolute numbers sheds light on how the phenomenon can arise: the distribution of cases across age groups differs significantly between the two countries, i.e., there is a statistical association between the country of reporting and the proportion of confirmed cases per age group. In particular, Italy recorded a much higher proportion of confirmed cases in older patients compared to China. This is illustrated in Fig. 1b (see Table 2 in Appendix A for exact numbers).

While most cases in China fell into the age range of 30–59, the majority of cases reported in Italy were in people aged 60 and over who are generally at higher risk of dying from Covid-19, as illustrated by the increase in CFRs with age shown in Fig. 1a for both countries. The observed difference may partly stem from the fact that the Italian population in general is older than the Chinese one with median ages of 45.4 and 38.4 respectively (see Table 3 in Appendix A for full age demographics of both countries), but additional factors such as different testing strategies and patterns in the social contacts among older and younger generations [e.g., 14] may also play a role. In summary, the larger share of confirmed cases among elderly people in Italy shown in Fig. 1b, combined with the fact that the elderly are generally at higher risk when contracting Covid-19, explains the mismatch between total and age-stratified CFRs shown in Fig. 1a and thus gives rise to Simpson's paradox in the data.[2]

3 A causal model for Covid-19 CFR data

While the previous reasoning provides a perfectly consistent explanation in a *statistical* sense, the phenomenon may still seem puzzling as it defies our *causal* intuition—similar to how an optical illusion defies our visual intuition. Humans appear to naturally extrapolate conditional probabilities to read them as causal effects, which can lead to inconsistent conclusions and may leave one wondering: *how can the disease in Italy be less fatal for the young, less fatal for the old, but more fatal for the people overall?* It is for this reason of ascribing causal meaning to probabilistic statements, that the reversal of (conditional) probabilities in §2 is perceived as and referred to as a "paradox" [15].

The aspiration to extrapolate causal conclusions from data is particularly strong in the context of a pandemic, during which many inherently causal questions are naturally asked. For example, politicians and citizens may want to evaluate different strategies to fight the disease by asking interventional (*"what if ...?"*) or counterfactual (*"what would have happened if ...?"*) questions. However, it is a well-known scientific mantra that "correlation does not imply causation", and observational data alone (like that in Fig. 1) is generally insufficient to draw causal conclusions. While correlations can be seen as a result of underlying causal mechanisms [20], different causal models can explain the same statistical association patterns equally well [17, 18]. Additional assumptions on the underlying causal structure are therefore necessary to guide reasoning based on observational data.

3.1 Assumptions

Included variables We consider the following three variables for comparing Covid-19 CFRs across different countries: (i) the *country* C in which a confirmed case is reported, modelled as a categorical variable; (ii) the *age group* A of a positively-tested patient, an ordinal variable with 10-year intervals as values; and (iii) the *medical outcome, or mortality,* M, a binary variable indicating whether a patient has deceased by the time of reporting ($M = 1$) or not ($M = 0$).

Data generating process and causal graph We assume the causal graph shown in Fig. 2, motivated by thinking of the following data-generating process: (1.) Choosing a country at random; (2.) Given the selected country, sampling a positively-tested patient with age group A; (3.) Conditional on the choice of C and A, sampling the mortality M. This is clearly a very simple and coarse-grained view of what is known to be a complex underlying phenomenon. As a consequence, we abstract away various influences and mechanisms within the arrows in Fig. 2. In particular, this view encompasses at least the following influences:

- ($C \to A$) encodes that the age distribution of cases is country-dependent. This difference might be due to a general difference in age demographic between countries, but other mechanisms such as inter-generational mixing or age-targeted social distancing may also play a role.
- ($A \to M$) reflects the notion that the disease is more dangerous for the elderly, i.e., age appears to have a causal effect on mortality.
- ($C \to M$) summarises country-specific influences on mortality other than age, e.g., approaches to testing, lockdown strategy and other non-pharmaceutical interventions, air pollution levels, and medical infrastructure, e.g., availability of hospital beds and ventilators. We will refer to the combination of all these effects as a country's *approach*.

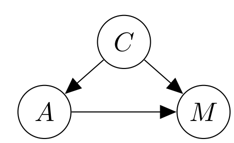


Figure 2: Assumed causal graph: note how, within this view, age A acts as a *mediator* of the effect of country C on mortality M.

Causal sufficiency and observational sample In addition, we assume causal sufficiency, meaning that all common causes of C, A, M are observed (i.e., there are no hidden confounders). Further, we assume that CFRs and the proportion of cases by age group are based on an observational sample and thus constitute estimates of $P(M = 1|A = a, C = c)$ and $P(A = a|C = c)$, respectively. We discuss these assumptions in the context of selection bias in §6.

[2] Note that the observed phenomenon can indeed *only* be explained if there is some association between the country and the number of confirmed cases per age group: if we simply took a weighted average of the CFRs shown in Figure 1a using the same weights for both countries, Simpson's paradox could not not arise since $a_i \leq b_i$ for $i = 1, ..., n$ implies that $\sum_{i=1}^n w_i a_i \leq \sum_{i=1}^n w_i b_i$ for any set of weights $\sum_{i=1}^n w_i = 1$.

4 Total, direct, and indirect (age-mediated) causal effects on mortality

Having clearly stated our assumptions, we can now answer causal queries within the model postulated in §3.1. In this section, we review definitions of different causal effects (following the treatment in [15]) and provide interpretations thereof by phrasing them as questions about different aspects of the CFR data in Fig. 1. We defer a discussion of issues such as identifiability under different conditions to Appendix B. Example calculations for each defined quantity using the data from Fig. 1 can be found in Appendix C. Throughout, we denote an intervention that externally fixed a variable X to a particular value x (as opposed to conditioning on it) using the notation $do(X = x)$ [16].

Total causal effect (TCE) First, we may ask about the overall causal effect of country on mortality:

Q_{TCE}: *"What would be the effect on mortality of changing country from China to Italy?"*

The answer to this query is called the average *total causal effect* (TCE):

Definition 1 (TCE). *The* TCE *of a binary treatment T on Y is defined as the interventional contrast*

$$\text{TCE}_{0\to1} = \mathbb{E}_{Y|do(T=1)}[Y|do(T = 1)] - \mathbb{E}_{Y|do(T=0)}[Y|do(T = 0)]. \tag{1}$$

In our setting (i.e., according to the causal graph in Fig. 2), the country C takes the role of a treatment that affects the outcome mortality M (denoted by T and Y, respectively, in Defn.1), and (subject to causal sufficiency) the TCE is simply given by the difference in total CFRs.

Asking *"why?"*: beyond total effects via mediation analysis While computing the TCE is the principled way to quantify the *total* causal influence, it does not help us understand what drives a difference between two countries, i.e., *why* it exists in the first place: we may also be interested in the *mechanisms* which give rise to different CFRs observed across different countries. Since the age of patients was crucial for explaining the instance of Simpson's paradox in §2, we now seek to better understand the role of age as a mediator of the effect of country on mortality. This seems particularly relevant from the perspective of countries, which—without being able to influence the age distribution of the general population—only have limited control over the age demographic of confirmed cases and thus may wish to factor out age-related effects.[3] However, such potential mediators are not reflected within the TCE, as evident from the absence of the age variable A from (9).

The country C causally influences mortality M along two different paths: a direct path $C \to M$, giving rise to a *direct effect*;[4] and an indirect path $C \to A \to M$ mediated by A, giving rise to an *indirect effect*. The TCE of C on M considered in §4 thus comprises both direct and indirect effects. Quantifying such direct and indirect effects is referred to as *mediation analysis* [15]. The main challenge is that any changes to the country C will propagate along both direct and indirect paths, making it difficult to isolate the different effects. The key idea is therefore to let changes propagate only along one path while somehow controlling or fixing the effect along the other path.

Controlled direct effect (CDE) The simplest way to measure a direct effect is by changing the treatment (country) while keeping the mediator fixed at a particular value. For example, we may ask about the causal effect for a particular age group such as 50–59 years olds:

$Q_{\text{CDE}(50-59)}$: *"For 50–59 year-olds, is it safer to get the disease in China or in Italy?"*

Because it involves actively *controlling* the value of the mediator, the answer to such a query is referred to as the average *controlled direct effect* (CDE). It is defined as follows.

Definition 2 (CDE). *The* CDE *of a binary treatment T on an outcome Y with mediator $X = x$ is*

$$\text{CDE}_{0\to1}(x) = \mathbb{E}[Y|do(T = 1, X = x)] - \mathbb{E}[Y|do(T = 0, X = x)]. \tag{2}$$

For our assumed setting, the CDE is given by the difference of CFRs for a given age group. A practical shortcoming of the CDE is that for real world scenarios it is often difficult or even impossible to control both the treatment and the mediator.[5] Another problem is that the CDE does not provide a global quantity for comparing baseline and treatment: in our setting, there is a different CDE *for each age group*. However, we may instead want to measure a direct effect at the *population level*.

[3] The demographic of confirmed cases can be influenced, e.g., via measures such as targeted isolation of the elderly, see also the discussion of the arrow $C \to A$ in §3.1.

[4] Recall that the direct effect of country on mortality is likely mediated by additional variables, which are subsumed in $C \to M$ in the current view—see §6 for further discussion.

[5] In medical settings, for example, one generally cannot easily control individual down-stream effects of a drug within the body, such as fixing, e.g., blood glucose levels while changing treatments.

Natural direct effect (NDE) Instead of fixing the mediator to a specific value (selecting a particular age group), we can consider the hypothetical question of what would happen under a change in treatment (country) if the mediator (age) kept behaving as it would under the control, i.e., as if the change only propagated along the direct path. This corresponds to asking about the effect of switching country without affecting the age distribution across the confirmed cases.

Q_{NDE}: *"For the Chinese case demographic, would the Italian approach have been better?"*

As it relies on the distribution of the mediator (age) under the control (China) to evaluate the treatment ("switching approach"), the answer to Q_{NDE} is known as average *natural direct effect* (NDE).

Definition 3 (NDE)**.** *The* NDE *of a binary treatment T on an outcome Y mediated by X is given by*

$$\text{NDE}_{0\to 1} = \mathbb{E}[Y_{X(0)}|do(T=1)] - \mathbb{E}[Y|do(T=0)]. \tag{3}$$

where $X(0)$ refers to the counterfactual distribution of X had T been 0.

Natural indirect effect (NIE) For isolating the indirect effect that a country exhibits on mortality only via age, $C \to A \to M$, we run into the additional complication that it is not possible to keep the influence along the direct path $C \to M$ constant under a change in treatment (country). To overcome this problem, one can consider a hypothetical change in the distribution of the mediator (age) as if the treatment (country) were changed, but without actually changing it. E.g., we may ask:

Q_{NIE}: *"How would the overall* CFR *in China change if the case demographic had instead been that from Italy, while keeping all else (i.e., the* CFR*'s of each age group) the same?"*

Since this considers a change of the mediator (age) to the natural distribution it would follow under a change treatment (case demographic from Italy) while keeping the treatment the same (Chinese CFR's), the answer to this question is referred to as the average *natural indirect effect* (NIE).

Definition 4 (NIE)**.** *The* NIE *of a binary treatment T on an outcome Y with mediator X is given by*

$$\text{NIE}_{0\to 1} = \mathbb{E}[Y_{X(1)}|do(T=0)] - \mathbb{E}[Y|do(T=0)]. \tag{4}$$

Mediation formulas: effects in causally sufficient systems For causally sufficient systems, the interventional distributions of each variable given its causal parents equal the corresponding observational distributions [16], corresponding to the intuition that they represent *mechanisms* rather than mere mathematical constructs [18]. This means that TCE (1) and CDE (2) reduce to:

$$\text{TCE}^{\text{obs}}_{0\to 1} = \mathbb{E}[Y|T=1] - \mathbb{E}[Y|T=0], \tag{5}$$
$$\text{CDE}^{\text{obs}}_{0\to 1}(x) = \mathbb{E}[Y|T=1, X=x] - \mathbb{E}[Y|T=0, X=x]. \tag{6}$$

Moreover, in this case, NDE (3) and NIE (4) are given by the following *mediation formulas* [15]:

$$\text{NDE}^{\text{obs}}_{0\to 1} = \textstyle\sum_x P\big(X=x|T=0\big)\big(\mathbb{E}[Y|T=1, X=x] - \mathbb{E}[Y|T=0, X=x]\big), \tag{7}$$
$$\text{NIE}^{\text{obs}}_{0\to 1} = \textstyle\sum_x \big(P(X=x|T=1) - P(X=x|T=0)\big)\mathbb{E}[Y|T=0, X=x]. \tag{8}$$

When comparing CFRs across countries, we only have access to observational data and thus rely on the assumption of causal sufficiency to compute total, direct and indirect effects via (5), (6) (7), (8).

Relation between TCE, NDE, and NIE *Can the total causal effect be decomposed into a sum of direct and indirect contributions?* While such an additive decomposition indeed exists for linear models,[6] it does not hold in general due to possible interactions between treatment and mediator, referred to as *moderation*.[7] Direct and indirect effects are not uniquely defined in general, but depend on the value of the mediator. Counterfactual quantities such as NDE and NIE are thus useful tools to measure some average form of direct and indirect effect with a meaningful interpretation.

[6]in which causal effects can be seen as path coefficients that can be multiplied to obtain path-specific effects

[7] [17] gives the illustrative example of a drug (treatment) that works by activating some proteins (mediator) inside the body before jointly attacking the disease: the drug is useless without the activated proteins (so the direct effect is zero) and the activated protein is useless without the chemical compound of the drug (so the indirect effect is also zero), but the total effect is non-zero because of the interaction between the two.

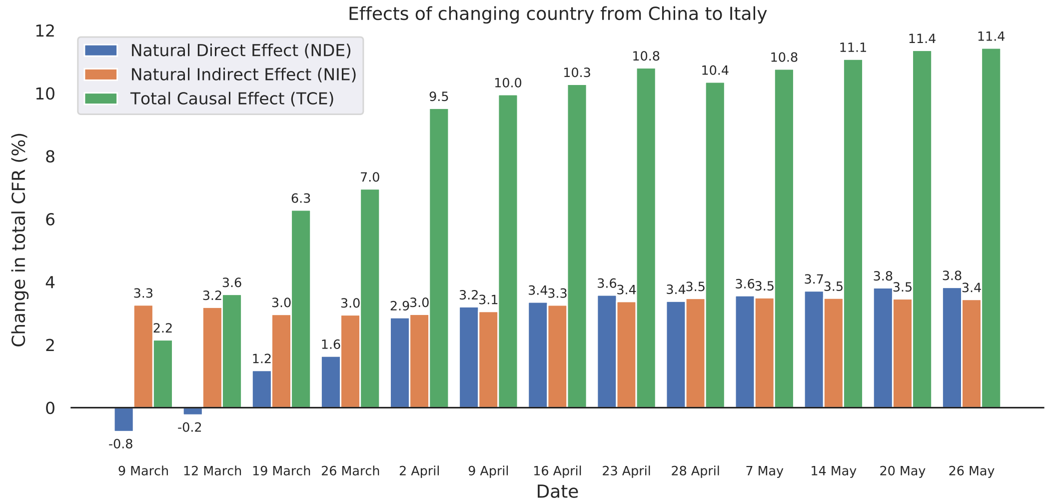


Figure 3: Evolution of TCE, NDE, and NIE of changing country from China to Italy on total CFR over time. We compare static data from China [27] with different snapshots from Italy reported by [10]. The direct effect initially was negative, meaning that age-specific mortality in Italy was lower; however, it changes sign around mid-March when an overloaded health system in northern Italy was reported [1]. The indirect effect remains mostly constant at a substantial +3–3.5%.

5 Case study: mediation analysis of age-related effects on Covid-19 CFRs

Dataset To employ the tools from mediation analysis outlined in §4 to better understand the influence of age on Covid-19 CFRs, we curated a dataset of confirmed cases and fatalities by age group (0–9, 10–19, etc.) from eleven countries (Argentina, China, Colombia, Italy, Netherlands, Portugal, South Africa, Spain, Sweden, Switzerland, South Korea) and the Diamond Princess cruise ship, on which the disease spread among passengers forced to quarantine on board [22]. The dataset includes 756,004 cases and 68,508 fatalities (total cumulative CFR of 9.06%), reported either by the different countries' national health institutes or in scientific publications. The selection of countries is based on availability of suitable data at the time of writing.[8] Where available, we included several reports from the same country, e.g., for Italy and Spain in weekly intervals. The data and our analysis (in form of an interactive notebook) are provided in the supplement and will be made publicly available. The exact sources and several additional figures and tables can be found in Appendices D and E.

Tracing causal effects over time First, we investigate the temporal evolution of direct and indirect (age-mediated) causal effects on mortality by expanding on the comparison from §2. The result of tracing TCE, NDE, and NIE of changing from China to Italy over a period of 11 weeks using (approximately) weekly reports from [10] is shown in Fig. 3. Note that case and fatality numbers for China remain constant in the figure, so that any changes over time can be attributed to Italy.[9]

We find that the TCE—which measures what would happen to the total CFR if *both* CFRs by age group *and* case demographic were changed to those from Italy—is positive throughout, reflecting a higher total CFR in Italy. It increases rapidly from an initial 2.2% (see Fig. 1a) to 9.5% over the first three weeks considered, and then continues to rise more slowly to 11.4%. This indicates that the difference between the two countries' total CFR becomes more pronounced over the time. In order to understand what drives this difference, we next consider the direct and indirect effects separately.

The NDE—which captures what would happen to the total CFR if the case demographic were kept the same, while only the approach (CFRs per age group) were changed— is negative at first, meaning that the considered change in approach would initially be beneficial, consistent with the lower CFRs in each age group shown in Fig. 1a. However, at a turning point around mid March the NDE changes

[8] Unfortunately, conventions on how to group patients by age vary across countries: e.g., Belgium, Canada, France, and Germany do not consistently use 10-year intervals; others such as the US use different groupings (0–4, 5–14, etc). For some countries (e.g., Brazil, Russia, Turkey, UK) we did not find demographic data.

[9] We did not find updated data beyond [27], though not many new cases have been reported from China since.

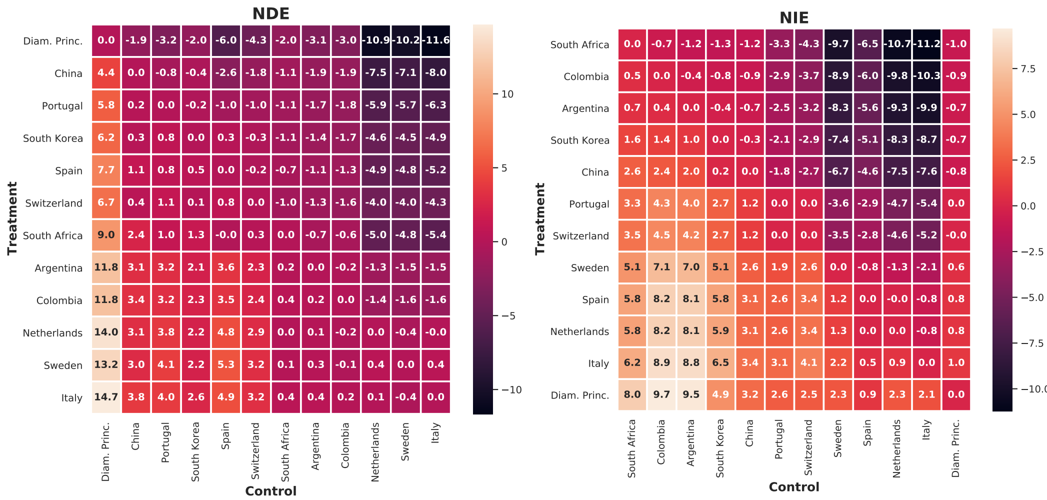


Figure 4: NDEs (left) and NIEs (right) for switching from the control country (columns) to the treatment country (rows). Numbers show the change in total CFR in %, i.e., negative numbers indicate that switching to the treatment country's approach, i.e., its CFRs by age group, (NDE) or case demographic (NIE) would lead to a decrease in total CFR. Countries are ordered by their average effect as a treatment country (NDE or NIE) over the remaining 11 data points as a control.

sign: beyond this point, switching to the Italian approach would lead to an increase in total CFR. While we can only speculate about the precise factors that came together in producing this reversal in NDE, it seems worth pointing out that a number of articles reported an overwhelmed health care system "close to collapse" in (northern) Italy during that very period of early to mid-March [1]. The NDE then keeps rising steeply until early April before gradually flattening off, similar to the TCE.

The NIE—which measures what would happen to total CFR if the approach were kept the same, while the case demographic were changed to that in Italy—on the other hand, remains largely constant over time, fluctuating between 3 and 3.5%, indicating that the case demographic in Italy does not change much over time. Its large value of over 3% means that simply changing the case demographic from China to that in Italy would already lead to a substantial increase in total CFR, consistent with the larger share of confirmed cases amongst the elderly in Italy shown in Fig. 1b.

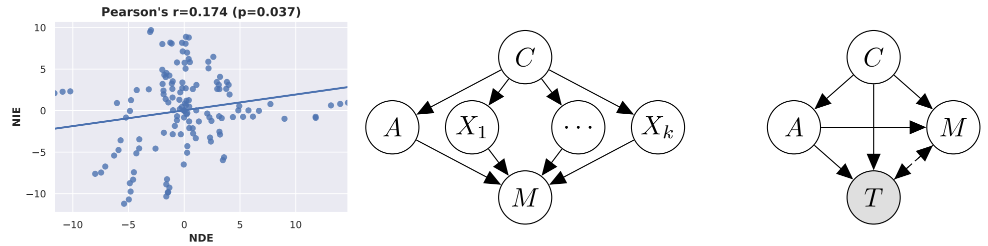
In summary, *while indirect age-related effects considerably contribute to differences in total* CFR—especially initially, when the instance of Simpson's paradox from §2 is reflected in the opposite signs of NDE and NIE—*it is mainly the direct effect that drives the observed changes over time.*

Comparison between several different countries We now leave the specific example of China vs Italy aside and turn to a comparison of different causal effects on Covid-19 mortality between the 12 countries (incl. the Diamond Princess) contained in our dataset.[10] All pairwise effects on total CFR (in %) of changing only *"approach"*, i.e., the CFRs by age group, (NDE; left) or case demographic (NIE; right) from a control country (columns) to a treatment country (rows) are shown in Fig. 4.

For ease of visualisation, the order in which countries are presented in Fig. 4 was chosen according to their average effect as a treatment over the remaining countries as control (i.e., by the mean of rows) for NDE and NIE separately. This allows to read off trends about the effectiveness of different approaches and the influence of the case demographic (subject to limitations such as, e.g., differences in testing which we discuss further in §6). In the case of NDE, for example, the Diamond Princess, China, Portugal, and South Korea compare favourably to most others in terms of their approaches, while the Netherlands, Sweden, and Italy occupy the bottom end of the range. In the case of NIE, on the other hand, South Africa, Colombia, and Argentina benefit most from their case demographic compared to other countries, while Spain, the Netherlands, Italy and the Diamond Princess are particularly affected by an adverse age distribution across confirmed cases.

Notably, there is no significant correlation between countries' ranking by NDE and NIE (Spearman's $\rho = 0.04$, $p = 0.9$), suggesting that a country's approach and case demographic may be largely

[10] We use the latest reported numbers available at the time of writing: except for China (17 February), the Diamond Princess (26 March), and Sweden (18 May) all reports are from the period of 25–29 May.



(a) (b) (c)

Figure 5: (a) NDE and NIE are only weakly correlated. (b) The direct effect $C \to M$ is likely mediated by additional variables $X_1, \ldots, X_k$. (c) Testing strategy may introduce selection bias, since CFR data implicitly conditions on having tested positive, represented by the shaded T.

unrelated. While some countries such as South Korea, Switzerland, the Netherlands, and Italy take almost the same place according to both rankings of particular interest are those countries for which rankings by NDE and NIE differ most. Other than for the Diamond Princess—which due to small sample size and high testing rates constitutes an illustrative special case that we discuss further in §6—the case of high ranking (rk) in terms of NDE and low ranking in terms of NIE is most most pronounced for Spain ($\text{rk}_{\text{NDE}} - \text{rk}_{\text{NIE}} = -4$), Portugal ($-3$), and China ($-3$). This suggests that, for the case of Spain, the high total CFR may, at least in parts, be attributed to an unfavourable case demographic, while the approaches (age-specific mortality) of China and Portugal may be even better than suggested by their (already comparatively low) total CFRs. Conversely, countries that rank considerably higher in terms of NIE than NDE include Colombia ($+7$), South Africa ($+6$), and Argentina ($+5$). These countries' low total CFRs may thus wrongly suggest a very successful approach while the low total CFR may actually, at least in parts, be due to an advantageous case demographic—again, subject to caveats such as differences in testing, see §6 for more details.

Noting that South Africa, Colombia, and Argentina are also the three youngest amongst the considered countries in terms of median age, we computed the Spearman correlation between the ranking of countries by NIE (as shown in Fig. 4) and by their median age and found a strong correlation between the two ($\rho = 0.94, p = 7 \times 10^{-6}$). This indicates that, for the countries considered, the case demographic is predominantly determined by the age distribution of the population, and suggests that countries seem not to make (effective) use of strategies such as, e.g., age-specific quarantines.

As a further investigation into the relation between direct and indirect effects on Covid-19 mortality, we find that, of the 132 ordered pairs of distinct countries, 64 exhibit opposite signs of NDE and NIE (as for the example of Simpson's paradox in §2, see also dates from early March in Fig. 3), meaning that *comparing countries in terms of total* CFR *may not give an accurate picture of the relative effectiveness of two countries' approaches* in those cases. Overall, pairwise NDEs and NIEs are only weakly but significantly correlated (Pearson's $r = 0.17, p = 0.04$) as shown in Fig. 5a.

6 Discussion

In this work, we have taken a coarse-grained causal modelling perspective considering the variables country C, age group A, and mortality M, which are commonly reported in the context of Covid-19 CFR data. This view abstracts away many potentially important factors (some of which we have named in §3.1) within the paths of the assumed causal graph. A strength of this approach is that it allows for consistent reasoning about age-mediated and non-age-related effects within the assumed model in situations where the data does not support a more fine-grained analysis. On the other hand, any conclusions drawn must be interpreted within this coarse-grained framework: we have thus referred to various country-specific influences on mortality collectively as "approach".

Considering additional mediators It is safe to assume that the virus is ultimately agnostic to the notion of different "*countries*" and that the influence of country on mortality $C \to M$ is not actually a direct one, but instead mediated by additional variables X_i, as illustrated in Fig. 5b. Candidates for such additional mediators X_i include, e.g., non-pharmaceutical interventions and critical healthcare infrastructure. We believe that many questions of interest regarding the Covid-19 pandemic can be phrased as path-specific causal effects involving such mediators, e.g.: *"What would be the effect on total* CFR *if country C_1 bought as many ventilators as country C_2?"*. Assuming more fine-grained data will become available as the pandemic progresses, extending our model with additional mediators and investigating their effects by building on the tools described in §4 is a promising future direction to deepen our understanding about which factors most drive Covid-19 mortality.

Testing strategy and selection bias An important potential limitation of our approach (or, more fundamentally, of CFR data) is that we only consider confirmed cases, i.e., patients who tested positively for Covid-19. We can make this explicit in our model by including test status T as additional variable. Our data is then always conditioned on $T = 1$, as illustrated in Fig. 5c. Since who is tested is not random, but generally depends both on a country's testing strategy and a patient's age (e.g., via severity of symptoms), reflected by the arrows $\{C, A\} \to M$ in Fig. 5c, this results in a problem of *selection bias* [19]. This issue is particularly clear for the case of the Diamond Princess on which *"3,063 PCR tests were performed among* [the 3,711] *passengers and crew members. Testing started among the elderly passengers, descending by age"* [22]. As a result of such extensive testing, the proportion of asymptomatic cases on board was very high (318 out of 619 detected cases), leading to low CFRs as manifested in the negative NDEs for the Diamond Princess as treatment in Fig. 4. This rate of testing is presently not feasible for countries with millions of inhabitants. Since testing capacities differ across countries, the reported CFR's may thus often not be comparable. Furthermore, a second source of selection bias may stem from the choice of countries which were included in our dataset: we only considered countries that report CFRs for age groups separately—those might be particularly affected by the pandemic. The cumulative CFR of 9% in our data may thus be inflated by such selection processes. Building on recent work in the causal inference literature on recoverability from selection bias may help account for this aspect of the problem [2, 5].

7 Conclusion

We have shown how causal reasoning can guide the interpretation of data of the ongoing pandemic. In particular, mediation analysis provides tools of separating effects due to different factors which, if not properly identified, can lead to misleading conclusions. We exploited these tools to uncover patterns in the time evolution of the pandemic in Italy, and in the comparison of multiple countries. In order to study age-mediated and age-unrelated effects on CFR across different countries, we curated a large-scale dataset from a multitude of sources; to the best of our knowledge, data allowing for this kind of analysis has not been aggregated before.

All datasets and an interactive notebook to reproduce and expand on our analyses are publicly available at: `github.com/Juliusvk/Covid19-age-related-causal-effects`

Acknowledgements

We are grateful to Elias Bareinboim for very detailed and insightful feedback on an earlier version of this manuscript.

References

[1] Benedetta Armocida, Beatrice Formenti, Silvia Ussai, Francesca Palestra, and Eduardo Missoni. The Italian health system and the COVID-19 challenge. *The Lancet Public Health*, 2020.

[2] Elias Bareinboim and Jin Tian. Recovering causal effects from selection bias. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[3] Centro de Coordinación de Alertas y Emergencias Sanitarias; (CCAES), Ministerio de Sanidad, Consumo y Bienestar Social; (MISAN, Ministry of Health, Consumer Affairs and Social Welfare). 29 May report. `https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/documentos/Actualizacion_120_COVID-19.pdf`, 2020.

[4] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCov and naming it SARS-CoV-2. *Nature Microbiology*, 2020.

[5] Juan Correa, Jin Tian, and Elias Bareinboim. Adjustment criteria for generalizing experimental findings. In *International Conference on Machine Learning*, pages 1361–1369, 2019.

[6] Federal Office of Public Health Switzerland. 26 May report. `https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html#-1199962081`, 2020.

[7] Folkhalsomyndigheten (Public Health Agency of Sweden). 18 May report. `https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Sweden#cite_note-306`, 2020. Accessed 29 May.

[8] Health Department republic of South Africa. 28 May report. `https://sacoronavirus.co.za/2020/05/29/update-on-covid-19-28th-may-2020/`, 2020. Accessed 29 May.

[9] Instituto Nacional de Salud. 28 May report. `https://www.ins.gov.co/Noticias/Paginas/Coronavirus.aspx`, 2020.

[10] Istituto Superiore di Sanità. Epidemia COVID-19: Aggiornamento nazionale, 09 Marzo 2020 – ore 16:00. `https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_09-marzo-2020.pdf`, 2020.

[11] Istituto Superiore di Sanità (ISS, Italian National Institute of Health). 26 May report. `https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_26-maggio-2020.pdf`, 2020.

[12] Korea Centers for Disease Control and Prevention. 25 May report. `https://www.cdc.go.kr/board/board.es?mid=a20501000000&bid=0015&list_no=367317&act=view`, 2020.

[13] Ministry of Health of Argentina. 28 May report. `https://www.argentina.gob.ar/salud/coronavirus-COVID-19/sala-situacion`, 2020. Accessed 29 May.

[14] Joël Mossong, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, 5 (3), 2008.

[15] Judea Pearl. Direct and indirect effects. In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*, pages 411–420, 2001.

[16] Judea Pearl. *Causality*. Cambridge University Press, 2009.

[17] Judea Pearl and Dana Mackenzie. *The book of why: the new science of cause and effect*. Basic Books, 2018.

[18] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms*. MIT Press, 2017.

[19] Dimple D Rajgor, Meng Har Lee, Sophia Archuleta, Natasha Bagdasarian, and Swee Chye Quek. The many estimates of the COVID-19 case fatality rate. *The Lancet Infectious Diseases*, 2020.

[20] Hans Reichenbach. *The Direction of Time*. University of California Press, Berkeley, CA, 1956.

[21] Rijksinstituut voor Volksgezondheid en Milieu. 28 May report. `https://www.rivm.nl/documenten/epidemiologische-situatie-covid-19-in-nederland-28-mei-2020`, 2020.

[22] Timothy W Russell, Joel Hellewell, Christopher I Jarvis, Kevin Van Zandvoort, Sam Abbott, Ruwan Ratnayake, Stefan Flasche, Rosalind M Eggo, W John Edmunds, Adam J Kucharski, et al. Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. *Eurosurveillance*, 25(12):2000256, 2020.

[23] Servico Nacional de Saude Republica Portuguesa. 28 May report. `https://covid19.min-saude.pt/wp-content/uploads/2020/05/87_DGS_boletim_20200528.pdf`, 2020.

[24] Edward H Simpson. The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 13(2):238–241, 1951.

[25] Johns Hopkins University. COVID-19 dashboard by the center for systems science and engineering (csse). `https://coronavirus.jhu.edu/map.html`, Retrieved 3 June 2020.

[26] WHO. Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV), 2020.

[27] Zunyou Wu and Jennifer M McGoogan. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *Jama*, 2020.

A Further material on the comparison China vs. Italy

In this Appendix, we provide additional details on the comparison of Italy and China that gives rise to the instance of Simpson's paradox in §2 and that was further investigated with a longitudinal approach in Fig. 3.

Tables 1, 2, and 3 show the CFRs, case demographic, and demographic of the general population for the two countries, respectively.

The relationship between case demographic and demographic of the general population is further investigated and visualised in Fig. 6.

Fig. 7 shows the temporal evolution of age-specific CFRs and case demographic for the longitudinal data from Italy used in Fig. 3.

Table 1: Exact numbers for the comparison of case fatality rates (CFRs) by age group for Italy and China shown in Fig. 1. Absolute numbers of fatalities/confirmed cases are shown in brackets below. Lower CFRs are highlighted in bold face. Sources: [27] and [10].

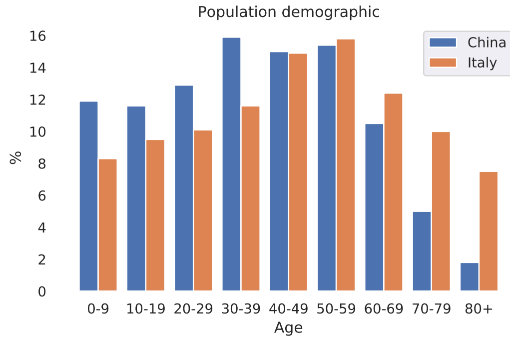
Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	$\geq$ 80	Total
Italy	**0%** (0/43)	**0%** (0/85)	**0%** (0/296)	**0%** (0/470)	**0.1%** (1/891)	**0.2%** (3/1,453)	**2.5%** (37/1,471)	**6.4%** (114/1,785)	**13.2%** (202/1,532)	4.4% (357/8,026)
China	**0%** (0/0)	0.2% (1/549)	0.2% (7/3,619)	0.2% (18/7,600)	0.4% (38/8,571)	1.3% (130/10,008)	3.6% (309/8,583)	8% (312/3,918)	14.8% (208/1,408)	**2.3%** (1,023/44,672)

Table 2: Proportion of confirmed cases from Table 1 by age group. This corresponds to the case demographics shown in Fig. 1b.

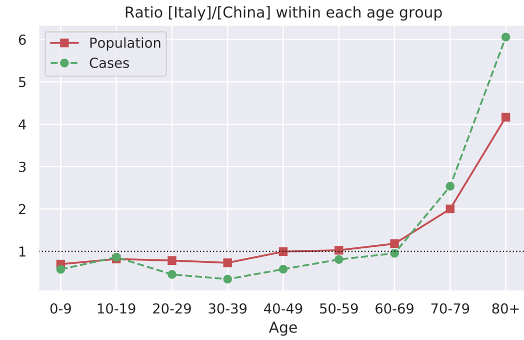
Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	$\geq$ 80
Italy	0.5%	1.0%	3.5%	5.6%	10.7%	17.4%	17.7%	**21.4%**	**18.4%**
China	**0.9%**	**1.2%**	**8.1%**	**17.0%**	**19.2%**	**22.4%**	**19.2%**	8.8%	3.2%

Table 3: Age demographic of the general population for Italy and China.

Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	$\geq$ 80
Italy	8.3%	9.5%	10.1%	11.6%	14.9%	**15.8%**	**12.4%**	**10%**	**7.5%**
China	**11.9%**	**11.6%**	**12.9%**	**15.9%**	**15%**	15.4%	10.5%	5%	1.8%



(a) Population demographic



(b) Ratios

Figure 6: Visualisation of the data from Tables 2 and 3 for the demographic comparison of China and Italy. (a) Demographic of the general population in the two countries (c.f. Fig. 1b). (b) Ratios (Italy / China) of the proportion of confirmed cases by age group (shown in dashed green) and the proportion of the general population within each age group from Table 3 (shown in solid red).

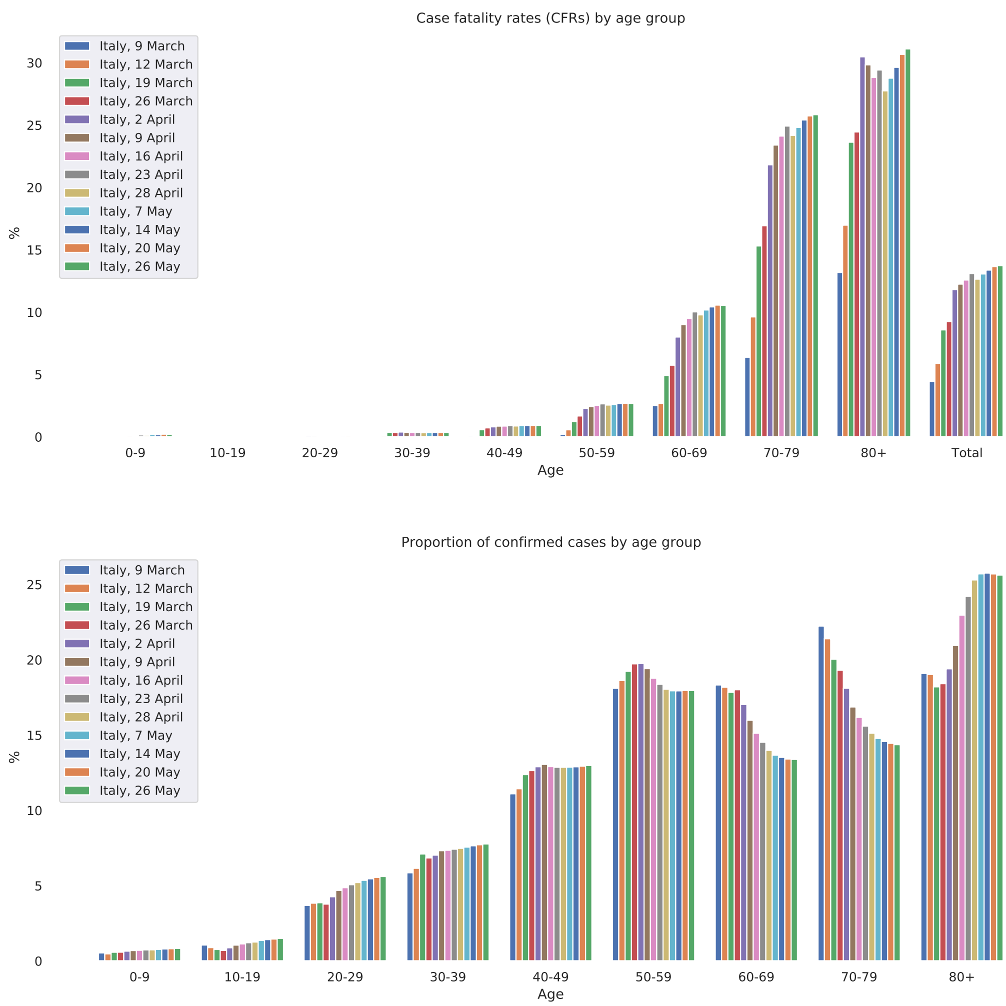


Figure 7: Different snapshots from Italy show the temporal evolution of CFRs by age group (top) and case demographic (bottom) over the time period for which different causal effects with China as a control country are shown in Fig. 3.

B Additional concepts from mediation analysis

B.1 Experimental (non-)identifiability of direct and indirect effects

Since the CDE in (2) only involves interventional quantities it is in principle *experimentally identifiable*, meaning that it can be determined through an experimental study in which both the treatment and the mediator are randomised, thus providing valid estimates of $P(Y|do(T=t, X=x))$.

In contrast, NDE and NIE are, *in general* (i.e., without further assumptions), *not experimentally identifiable* owing to their counterfactual nature. However, under certain conditions such non-confoundedness of mediator and outcome experimental identifiability is obtained.[11] In this case:

$$\text{NDE}^{\text{exp}}_{0\to1} = \textstyle\sum_x P\big(X=x|do(T=0)\big)\left(\mathbb{E}[Y|do(T=1,X=x)] - \mathbb{E}[Y|do(T=0,X=x)]\right),$$
$$\text{NIE}^{\text{exp}}_{0\to1} = \textstyle\sum_x \big(P\big(X=x|do(T=1)\big) - P\big(X=x|do(T=0)\big)\big)\,\mathbb{E}[Y|do(T=0,X=x)]\,.$$

Note that even then, identifying natural effects requires combining results from two different experimental settings: one where both mediator and treatment are randomised, and a second in which treatment is randomised and the mediator observed. This again highlights the hypothetical nature of NDE and NIE and explains why they—unlike TCE and CDE—cannot simply be read off from a table like Table 1, even when causal sufficiency is assumed.

B.2 Subtractivity principle

There exists a general formula relating TCE, NDE, and NIE known as the *subtractivity principle* that follows from their definitions and holds without restrictions on the type of model [15]:

$$\text{TCE}_{0\to1} = \text{NDE}_{0\to1} - \text{NIE}_{1\to0} = \text{NIE}_{0\to1} - \text{NDE}_{1\to0}.$$

C Example calculations for TCE, CDE, NDE and NIE

C.1 TCE

To address Q_{TCE} in our example we need to compute

$$\text{TCE}_{\text{China}\to\text{Italy}} = \mathbb{E}[M|do(C=\text{Italy})] - \mathbb{E}[M|do(C=\text{China})]. \tag{9}$$

From the assumed causal graph and causal sufficiency, it follows that for our setting $P(A|do(C)) = P(A|C)$ and $P(M|do(A,C)) = P(M|A,C)$. We can thus compute (9) as

$$\begin{aligned}\text{TCE}_{\text{China}\to\text{Italy}} &= \sum_a \left[P_{M|A,C}(1|a,\text{Italy})P_{A|C}(a|\text{Italy}) - P_{M|A,C}(1|a,\text{China})P_{A|C}(a|\text{China})\right]\\ &\approx 2.2\%.\end{aligned}$$

Note that this corresponds to the difference of total CFRs reported in the last column of Table 1. This means that the difference of total CFRs indeed constitutes a causal effect, and changing country from China to Italy would lead to an overall increase in CFR of $\approx 2.2\%$ (given the data in Table 1 and subject to our modelling assumptions).

C.2 CDE

To address $Q_{\text{CDE}(a)}$ in our example, we need to compute

$$\begin{aligned}\text{CDE}_{\text{China}\to\text{Italy}}(a) &= \mathbb{E}[M|do(C=\text{Italy},A=a)] - \mathbb{E}[M|do(C=\text{China},A=a)]\\ &= P(M=1|do(C=\text{Italy},A=a)) - P(M=1|do(C=\text{China},A=a))\\ &= P(M=1|C=\text{Italy},A=a) - P(M=1|C=\text{China},A=a).\end{aligned}$$

This corresponds to the difference between CFRs across the two countries within a particular age group, i.e., the difference of two CFRs within a particular column of Table 1. Hence, the answer to $Q_{\text{CDE(50–59)}}$ is that for this age group it is safer to switch country to Italy with a resulting change in CFR of $\approx 0.2\% - 1.3\% = -1.1\%$. (Bear in mind that this calculation is based on Italian data from beginning of March.)

[11] A more general criterion is the existence of a set of covariates W, non-descendants of T and X, which satisfy the graphical d-separation criterion $(Y \perp\!\!\!\perp X|W)_{G_{\underline{\text{TX}}}}$, see [15, Thms. 1&4] for details.

C.3 NDE

Applying our assumptions, in particular causal sufficiency, we can calculate the NDE to answer Q_{NDE} for our running example as follows,

$$\begin{aligned}
\text{NDE}_{\text{China}\to\text{Italy}} &= \mathbb{E}[M_{A(\text{China})}|do(C = \text{Italy})] - \mathbb{E}[M_{A(\text{China})}|do(C = \text{China})] \\
&= \sum_a P_{A|do(C)}(a|do(\text{China}))\big[P_{M|do(A,C)}(1|do(a, \text{Italy})) - P_{M|do(A,C)}(1|do(a, \text{China}))\big] \\
&= \sum_a P_{A|C}(a|\text{China})\big[P_{M|A,C}(1|a, \text{Italy}) - P_{M|A,C}(1|a, \text{China})\big] \\
&= \mathbb{E}_{A|C=\text{China}}\big[\text{CDE}_{\text{China}\to\text{Italy}}(A)\big] \approx -0.8\%.
\end{aligned}$$

We thus find that when we only consider the Chinese case demographic, using the Italian approach (i.e., the CFRs for Italy from Table 1) would lead to a reduction in total CFR of $\approx 0.8\%$, consistent with our observation from §2 that CFRs were lower in Italy for each age group.

Remark 1. *As is apparent from the last line of the above calculation, the* NDE *can be interpreted as an expected* CDE *w.r.t. a particular (counterfactual) distribution of the mediator. Here, due to our assumption of causal sufficiency the expectation is taken w.r.t. the conditional distribution of A in the control group (China).*

Remark 2. *Taking the previous remark about* NDE *as the expected* CDE *within the control group one step further, we can, of course, also consider expected* CDE*s w.r.t. other distributions describing a target-population we want to reason about. For example, a third country, say Spain, may be considering whether to adopt the Chinese or Italian approach given its own case demographic. In this case, we would be interested in the following quantity.*

$$\mathbb{E}_{A|C=\textit{Spain}}[\text{CDE}_{\textit{China}\to\textit{Italy}}(A)] = \textstyle\sum_a P_{A|C}(a|\textit{Spain})\text{CDE}_{\textit{China}\to\textit{Italy}}(a)$$

C.4 NIE

Again, using causal sufficiency, we can calculate the NIE to answer Q_{NIE} for our example as follows,

$$\begin{aligned}
\text{NIE}_{\text{China}\to\text{Italy}} &= \mathbb{E}[M_{A=A_{\text{Italy}}}|do(C = \text{China})] - \mathbb{E}[M_{A=A_{\text{China}}}|do(C = \text{China})] \\
&= \sum_a \big[P_{A|do(C)}(a|do(\text{Italy})) - P_{A|do(C)}(a|do(\text{China}))\big] P_{M|do(A,C)}(1|do(a, \text{China})) \\
&= \sum_a \big[P_{A|C}(a|\text{Italy}) - P_{A|C}(a|\text{China})\big] P_{M|A,C}(1|a, \text{China}) \\
&\approx 3.3\%
\end{aligned}$$

We thus find that changing only the case demographic to that from Italy would lead to a substantial increase in total CFR in China of about 3.3%. Notably, the NIE is of the opposite sign of the NDE suggesting that indirect and direct effects are counteracting in our example as the reader may have expected from §2: despite the lower CFRs in each age group (leading to a negative NDE) the total CFR is larger in Italy due to the higher age of positively-tested patients (leading to a positive NIE).

C.5 Substractivity-principle

In our running example we find that

$$\text{TCE}_{\text{China}\to\text{Italy}} = 2.2\% \neq -0.8\% + 3.3\% = \text{NDE}_{\text{China}\to\text{Italy}} + \text{NIE}_{\text{China}\to\text{Italy}}$$

indicating that some level of moderation or interaction is present.

D Dataset details

In this Appendix, we provide further details on the datasets of age-stratified case and fatality numbers curated as part of this work. We provide three different datasets:

- A dataset containing the latest age-stratified case and fatality numbers for all different countries considered in our analysis, described in more detail in D.1.
- A dataset containing longitudinal age-stratified case and fatality numbers for Italy, described in more detail in D.2.
- A dataset containing longitudinal age-stratified case and fatality numbers for Spain, described in more detail in D.3.

All datasets are contained in the supplementary material in multiple commonly-used formats (`.csv`, `.xlsx`, `.json`, `.pkl`) and will be made publicly available upon publication.

D.1 Dataset of latest age-stratified case and fatality numbers for different countries

An overview of the dataset of latest age-stratified case and fatality numbers for different countries in the form of metadata is shown in Table 4.

CFRs and absolute case and fatality numbers in age-stratified form are shown in Table 5.

Case demographics are shown in Table 6.

Table 4: Information on the sources for the data regarding the countries in the case study.

Country	Date of reporting	Confirmed cases	Fatalities	Source
Argentina	28 May	14,675	507	[13]
China	17 February	44,672	1023	[27]
Colombia	28 May	25,366	822	[9]
Diam. Princ.	26 March	619	7	[22]
Italy	26 May	230,760	31676	[11]
Netherlands	28 May	45,947	5903	[21]
Portugal	28 May	31,596	1369	[23]
South Africa	28 May	27,280	577	[8]
South Korea	25 May	11,190	266	[12]
Spain	29 May	258,760	20585	[3]
Sweden	18 May	34,432	4125	[7]
Switzerland	26 May	30,707	1648	[6]

Table 5: Exact numbers for the comparison of case fatality rates (CFRs) by age group for all countries discussed in §5. Absolute numbers of fatalities/confirmed cases are shown in brackets below.

Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	≥ 80	Total
Argentina	0.0% (0/1,002)	0.1% (1/1,080)	0.0% (1/2,813)	0.3% (9/3,142)	1.0% (24/2,508)	3.0% (54/1,812)	10.0% (101/1,005)	18.9% (123/651)	29.3% (194/662)	3.5% (507/14,675)
China	0.0% (0/416)	0.2% (1/549)	0.2% (7/3,619)	0.2% (18/7,600)	0.4% (38/8,571)	1.3% (130/10,008)	3.6% (309/8,583)	8.0% (312/3,918)	14.8% (208/1,408)	2.3% (1,023/44,672)
Colombia	0.5% (5/1,105)	0.1% (1/1,950)	0.2% (13/5,614)	0.4% (24/5,615)	1.5% (61/4,033)	3.7% (121/3,286)	9.8% (192/1,961)	19.2% (214/1,117)	27.9% (191/685)	3.2% (822/25,366)
Diam. Princ.	0.0% (0/1)	0.0% (0/5)	0.0% (0/28)	0.0% (0/34)	0.0% (0/27)	0.0% (0/59)	0.0% (0/177)	1.3% (3/234)	7.4% (4/54)	1.1% (7/619)
Italy	0.2% (4/1,919)	0.0% (0/3,442)	0.1% (12/12,933)	0.3% (62/17,934)	0.9% (273/29,942)	2.7% (1,109/41,435)	10.6% (3,259/30,880)	25.8% (8,562/33,141)	31.1% (18,395/59,134)	13.7% (31,676/230,760)
Netherlands	0.0% (0/128)	0.2% (1/587)	0.1% (3/4,336)	0.2% (10/4,093)	0.5% (28/5,269)	1.7% (142/8,437)	8.1% (484/5,949)	25.6% (1,596/6,229)	33.3% (3,639/10,919)	12.8% (5,903/45,947)
Portugal	0.0% (0/626)	0.0% (0/1,052)	0.0% (1/4,114)	0.0% (1/4,736)	0.3% (15/5,315)	0.8% (42/5,253)	3.5% (122/3,484)	10.6% (269/2,537)	20.5% (919/4,479)	4.3% (1,369/31,596)
South Africa	0.3% (2/755)	0.1% (1/1,147)	0.1% (4/5,319)	0.4% (33/7,720)	1.1% (61/5,754)	3.8% (144/3,753)	9.2% (153/1,663)	15.0% (113/754)	15.9% (66/415)	2.1% (577/27,280)
South Korea	0.0% (0/149)	0.0% (0/636)	0.0% (0/3,117)	0.2% (2/1,235)	0.2% (3/1,481)	0.8% (15/1,987)	2.8% (39/1,375)	10.8% (78/719)	26.3% (129/491)	2.4% (266/11,190)
Spain	0.3% (3/1,123)	0.2% (5/2,068)	0.2% (24/15272)	0.3% (65/24,902)	0.6% (218/37,970)	1.4% (663/45750)	5.0% (1,825/36,355)	14.3% (4,896/34,294)	21.1% (12,886/61,026)	8.0% (20,585/258,760)
Sweden	0.6% (1/168)	0.0% (0/401)	0.3% (8/3,104)	0.3% (12/4,051)	0.8% (39/4,962)	2.1% (129/6,190)	7.0% (294/4,186)	23.4% (909/3,888)	36.5% (2,733/7,482)	12.0% (4,125/34,432)
Switzerland	0.0% (0/162)	0.0% (0/877)	0.0% (0/3,844)	0.1% (5/4,136)	0.1% (4/4,809)	0.6% (37/6,232)	3.3% (121/3,671)	11.6% (335/2,896)	28.1% (1,146/4,080)	5.4% (1,648/30,707)

Table 6: Proportion of confirmed cases by age group for all of the countries considered in section 5.

Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	≥ 80
Argentina	6.8%	7.4%	19.2%	21.4%	17.1%	12.3%	6.8%	4.4%	4.6%
China	0.9%	1.2%	8.1%	17.0%	19.2%	22.4%	19.2%	8.8%	3.2%
Colombia	4.4%	7.7%	22.1%	22.1%	15.9%	13.0%	7.7%	4.4%	2.7%
Diam. Princ.	0.2%	0.8%	4.5%	5.5%	4.4%	9.5%	28.6%	37.8%	8.7%
Italy	0.8%	1.5%	5.6%	7.8%	13.0%	18.0%	13.4%	14.4%	25.5%
Netherlands	0.3%	1.3%	9.4%	8.9%	11.5%	18.4%	12.9%	13.6%	23.7%
Portugal	2.0%	3.3%	13.0%	15.0%	16.8%	16.6%	11.0%	8.0%	14.3%
South Africa	2.8%	4.2%	19.5%	28.3%	21.1%	13.8%	6.1%	2.8%	1.4%
South Korea	1.3%	5.7%	27.9%	11.0%	13.2%	17.8%	12.3%	6.4%	4.4%
Spain	0.4%	0.8%	5.9%	9.6%	14.7%	17.7%	14.0%	13.3%	23.6%
Sweden	0.5%	1.2%	9.0%	11.8%	14.4%	18.0%	12.2%	11.3%	21.7%
Switzerland	0.5%	2.9%	12.5%	13.5%	15.7%	20.3%	12.0%	9.4%	13.2%

D.2 Dataset of longitudinal age-stratified case and fatality numbers for Italy

An overview of the dataset of longitudinal age-stratified case and fatality numbers for Italy, in the form of metadata, is shown in Table 7. The source for all different time points is the same as that shown in Table 4 for Italy, queried at the corresponding dates shown in Table 7.

CFRs and absolute case and fatality numbers in age-stratified form are shown in Table 8.

Case demographics are shown in Table 9.

Table 7: Metadata for the longitudinal data from Italy.

Date of reporting	Confirmed cases	Fatalities
9 March	8,026	357
12 March	13,317	785
19 March	35,529	3,047
23 March	57,695	5,018
26 March	73,534	6,801
2 April	106,231	12,548
9 April	135,968	16,653
16 April	159,003	19,994
23 April	177,025	23,118
28 April	199,389	25,215
7 May	214,047	27,955
14 May	222,022	29,691
20 May	227,153	31,017
26 May	230,760	31,676

Table 8: Age-specific CFRs for the longitudinal data for Italy.

Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	≥ 80	Total
9 March	0.0% (0/43)	0.0% (0/85)	0.0% (0/296)	0.0% (0/470)	0.1% (1/891)	0.2% (3/1453)	2.5% (37/1471)	6.4% (114/1785)	13.2% (202/1532)	4.4% (357/8026)
12 March	0.0% (0/63)	0.0% (0/118)	0.0% (0/511)	0.1% (1/819)	0.1% (1/1523)	0.6% (14/2480)	2.7% (65/2421)	9.6% (274/2849)	17.0% (430/2533)	5.9% (785/13317)
19 March	0.0% (0/205)	0.0% (0/270)	0.0% (0/1374)	0.4% (9/2525)	0.6% (25/4396)	1.2% (83/6834)	4.9% (312/6337)	15.3% (1090/7121)	23.6% (1528/6467)	8.6% (3047/35529)
23 March	0.0% (0/318)	0.0% (0/386)	0.0% (0/2192)	0.3% (12/3995)	0.6% (41/7267)	1.5% (168/11280)	5.2% (541/10423)	15.6% (1768/11320)	23.7% (2488/10514)	8.7% (5018/57695)
26 March	0.0% (0/428)	0.0% (0/512)	0.0% (0/2778)	0.3% (17/5033)	0.7% (67/9295)	1.7% (243/14508)	5.7% (761/13243)	16.9% (2403/14198)	24.4% (3310/13539)	9.2% (6801/73534)
2 April	0.0% (0/693)	0.0% (0/931)	0.1% (6/4530)	0.4% (29/7466)	0.8% (110/13701)	2.3% (479/20975)	8.0% (1448/18089)	21.8% (4196/19238)	30.5% (6280/20608)	11.8% (12548/106231)
9 April	0.1% (1/938)	0.0% (0/1432)	0.1% (7/6360)	0.4% (36/9956)	0.9% (153/17745)	2.4% (638/26391)	9.0% (1957/21734)	23.4% (5366/22934)	29.8% (8495/28478)	12.2% (16653/135968)
16 April	0.1% (1/1123)	0.0% (0/1804)	0.1% (7/7737)	0.3% (40/11686)	0.9% (178/20519)	2.5% (756/29858)	9.5% (2284/24040)	24.1% (6203/25717)	28.8% (10525/36519)	12.6% (19994/159003)
23 April	0.2% (2/1304)	0.0% (0/2146)	0.1% (7/8963)	0.4% (48/13137)	0.9% (203/22767)	2.6% (861/32524)	10.0% (2576/25707)	24.9% (6882/27615)	29.4% (12609/42862)	13.1% (23188/177025)
28 April	0.1% (2/1478)	0.0% (0/2511)	0.1% (8/10377)	0.3% (49/14907)	0.9% (224/25644)	2.6% (918/35986)	9.8% (2727/27880)	24.2% (7291/30158)	27.7% (13996/50448)	12.6% (25215/199389)
7 May	0.2% (3/1642)	0.0% (0/2908)	0.1% (9/11457)	0.3% (54/16189)	0.9% (246/27553)	2.6% (993/38399)	10.2% (2976/29252)	24.8% (7849/31627)	28.8% (15825/55020)	13.1% (27955/214047)
14 May	0.2% (3/1774)	0.0% (0/3148)	0.1% (12/12115)	0.3% (59/16981)	0.9% (258/28627)	2.7% (1063/39822)	10.4% (3127/30010)	25.4% (8221/32353)	29.6% (16948/57192)	13.4% (29691/222022)
20 May	0.2% (4/1851)	0.0% (0/3312)	0.1% (14/12599)	0.3% (61/17528)	0.9% (268/29390)	2.7% (1101/40803)	10.6% (3219/30466)	25.7% (8447/32824)	30.7% (17903/58380)	13.7% (31017/227153)
26 May	0.2% (4/1919)	0.0% (0/3442)	0.1% (12/12933)	0.3% (62/17934)	0.9% (273/29942)	2.7% (1109/41435)	10.6% (3259/30880)	25.8% (8562/33141)	31.1% (18395/59134)	13.7% (31676/230760)

Table 9: Proportion of confirmed cases by age group for longitudinal data for Italy.

Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	≥ 80
9 March	0.5%	1.1%	3.7%	5.9%	11.1%	18.1%	18.3%	22.2%	19.1%
12 March	0.5%	0.9%	3.8%	6.2%	11.4%	18.6%	18.2%	21.4%	19.0%
19 March	0.6%	0.8%	3.9%	7.1%	12.4%	19.2%	17.8%	20.0%	18.2%
23 March	0.6%	0.7%	3.8%	6.9%	12.6%	19.6%	18.1%	19.6%	18.1%
26 March	0.6%	0.7%	3.8%	6.8%	12.6%	19.7%	18.0%	19.3%	18.5%
2 April	0.7%	0.9%	4.3%	7.0%	12.9%	19.7%	17.0%	18.1%	19.4%
9 April	0.7%	1.1%	4.7%	7.3%	13.1%	19.4%	16.0%	16.9%	20.8%
16 April	0.7%	1.1%	4.9%	7.3%	12.9%	18.8%	15.1%	16.2%	23.0%
23 April	0.7%	1.2%	5.1%	7.4%	12.9%	18.4%	14.5%	15.6%	24.2%
28 April	0.7%	1.3%	5.2%	7.5%	12.9%	18.0%	14.0%	15.1%	25.3%
7 May	0.8%	1.4%	5.4%	7.6%	12.9%	17.9%	13.7%	14.7%	25.6%
14 May	0.8%	1.4%	5.5%	7.6%	12.9%	17.9%	13.5%	14.6%	25.8%
20 May	0.8%	1.5%	5.5%	7.7%	12.9%	18.0%	13.4%	14.5%	25.7%
26 May	0.8%	1.5%	5.6%	7.8%	13.0%	18.0%	13.4%	14.4%	25.5%

D.3 Dataset of longitudinal age-stratified case and fatality numbers for Spain

An overview of the dataset of longitudinal age-stratified case and fatality numbers for Spain, in the form of metadata, is shown in Table 10. The source for all different time points is the same as that shown in Table 4 for Spain, queried at the corresponding dates shown in Table 10.

CFRs and absolute case and fatality numbers in age-stratified form are shown in Table 11.

Case demographics are shown in Table 12.

Table 10: Metadata for the longitudinal data from Spain.

Date of reporting	Confirmed cases	Fatalities
22 March	18,959	805
26 March	32,816	1,326
30 March	51,626	2,784
2 April	69,177	4,361
9 April	106,447	6,729
16 April	133,082	10,793
23 April	152,687	13,078
28 April	204,866	15,853
7 May	220,444	17,460
14 May	239,095	19,115
29 May	258,760	20,585

Table 11: Longitudinal age-stratified data for Spain.

Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	≥ 80	Total
22 March	0.0% (0/129)	0.5% (1/221)	0.3% (4/1285)	0.1% (3/2208)	0.3% (9/2919)	0.6% (20/3129)	2.2% (63/2916)	5.2% (164/3132)	17.9% (541/3020)	4.2% (805/18959)
26 March	0.0% (0/175)	0.3% (1/302)	0.2% (4/1932)	0.2% (7/3454)	0.4% (19/5045)	0.6% (35/5749)	2.1% (114/5397)	5.6% (303/5377)	15.7% (843/5385)	4.0% (1326/32816)
30 March	0.0% (0/212)	0.3% (1/368)	0.2% (6/2883)	0.2% (10/5351)	0.5% (36/7965)	0.8% (78/9390)	2.7% (232/8744)	8.8% (759/8625)	20.5% (1662/8088)	5.4% (2784/51626)
2 April	0.0% (0/250)	0.2% (1/434)	0.2% (6/3590)	0.3% (18/6853)	0.5% (49/10551)	1.0% (131/12722)	3.2% (373/11657)	10.3% (1176/11368)	22.2% (2607/11752)	6.3% (4361/69177)
9 April	0.4% (1/285)	0.2% (1/588)	0.2% (11/5381)	0.2% (24/10341)	0.4% (61/16088)	1.0% (197/19836)	3.4% (597/17713)	10.5% (1773/16957)	21.1% (4064/19258)	6.3% (6729/106447)
16 April	0.2% (1/423)	0.3% (2/734)	0.3% (19/6763)	0.3% (37/12466)	0.6% (116/19536)	1.3% (312/24471)	4.5% (958/21249)	14.1% (2868/20287)	23.9% (6480/27153)	8.1% (10793/133082)
23 April	0.4% (2/502)	0.3% (3/869)	0.3% (25/7962)	0.3% (50/14304)	0.6% (138/22430)	1.4% (400/27795)	4.9% (1149/23595)	15.0% (3374/22470)	24.2% (7937/32760)	8.6% (13078/152687)
28 April	0.3% (2/660)	0.3% (4/1206)	0.2% (22/11138)	0.3% (55/18924)	0.6% (172/29629)	1.4% (497/36423)	4.6% (1387/30361)	13.6% (4012/29550)	20.7% (9702/46975)	7.7% (15853/204866)
7 May	0.3% (2/765)	0.4% (5/1398)	0.2% (21/12321)	0.3% (57/20759)	0.6% (185/32239)	1.4% (569/39418)	4.8% (1541/32226)	14.0% (4320/30861)	21.3% (10760/50457)	7.9% (17460/220444)
14 May	0.2% (2/871)	0.3% (5/1619)	0.2% (23/13439)	0.3% (62/22643)	0.6% (201/35175)	1.4% (610/42874)	4.9% (1693/34380)	14.3% (4628/32395)	21.4% (11931/55699)	8.0% (19155/239095)
29 May	0.3% (3/1123)	0.2% (5/2068)	0.2% (24/15272)	0.3% (65/24902)	0.6% (218/37970)	1.4% (663/45750)	5.0% (1825/36355)	14.3% (4896/34294)	21.1% (12886/61026)	8.0% (20585/258760)

Table 12: Proportion of confirmed cases by age group for the longitudinal data from Spain.

Age	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	≥ 80
22 March	0.5%	2.9%	12.5%	13.5%	15.7%	20.3%	12.0%	9.4%	13.2%
26 March	0.5%	0.9%	5.9%	10.5%	15.4%	17.5%	16.4%	16.4%	16.5%
30 March	0.4%	0.7%	5.6%	10.4%	15.4%	18.2%	16.9%	16.7%	15.7%
2 April	0.4%	0.6%	5.2%	9.9%	15.3%	18.4%	16.9%	16.4%	16.9%
9 April	0.3%	0.6%	5.1%	9.7%	15.1%	18.6%	16.6%	15.9%	18.1%
16 April	0.3%	0.6%	5.1%	9.4%	14.7%	18.4%	16.0%	15.2%	20.3%
23 April	0.3%	0.6%	5.2%	9.4%	14.7%	18.2%	15.5%	14.7%	21.4%
28 April	0.3%	0.6%	5.4%	9.2%	14.5%	17.8%	14.8%	14.4%	23.0%
7 May	0.3%	0.6%	5.6%	9.4%	14.6%	17.9%	14.6%	14.0%	23.0%
14 May	0.4%	0.7%	5.6%	9.5%	14.7%	17.9%	14.4%	13.5%	23.3%
29 May	0.4%	0.8%	5.9%	9.6%	14.7%	17.7%	14.0%	13.3%	23.6%

E Additional results and figures

E.1 Temporal CFR data for Spain

We perform a similar analysis of the temporal evolution of different causal effects of changing country from China to Spain, as done for Italy in §5 and Fig. 3. The results are shown in Fig. 8. Recall that the control China remains fixed throughout so that any changes can be attributed to changes in the Spanish data.

Interestingly, a reversal in the sign of the NDE can also be observed for Spain, taking place around 30 March. This bears similarity to the reversal of NDE observed for Italy. The initial increase in NDE is also reflected in the age-specific CFRs shown in the middle of Fig. 8 which are initally increasing for most age groups. Unlike Italy, however, NDE and TCE do not increase monotonically, but reach a maximum (over the time period considered) around 23 April and subsequently decrease again. The NIE also appears less constant than for the case of changing country to Italy shown in Fig. 3, steadily climbing from intially 2.3% to 3.1% at the end of May (ca. 35% increase).

As a remark of caution, we point out that the total number of fatalities reported by the Spanish ministry in age-stratified form is considerably lower than the number of fatalities reported (without separation into age groups) by different sources such as, e.g., [25]. This may have different reasons such as, e.g., latency in their reporting of fatalities in general, or of the exact age group of deceased patients specifically. As a result, CFRs from Spain are lower than other sources suggest, and may thus not be very reliable.

E.2 Comparison of age-specific CFRs and case demographic between different countries

A visual comparison of CFRs by age group and case demographic (similar to that shown in Fig. 1 for only China and Italy) for all different countries in our dataset is shown in Fig. 9.

E.3 TCEs between different countries

In addition to the pair-wise NDEs and NIEs between the different countries in our dataset, we also show the pair-wise TCEs for completeness in Fig. 10. Note that—as opposed to NDE and NIE—the TCE is, by definition, symmetric, i.e., $\text{TCE}_{0\to 1} = -\text{TCE}_{1\to 0}$, as can be seen from Fig. 10.

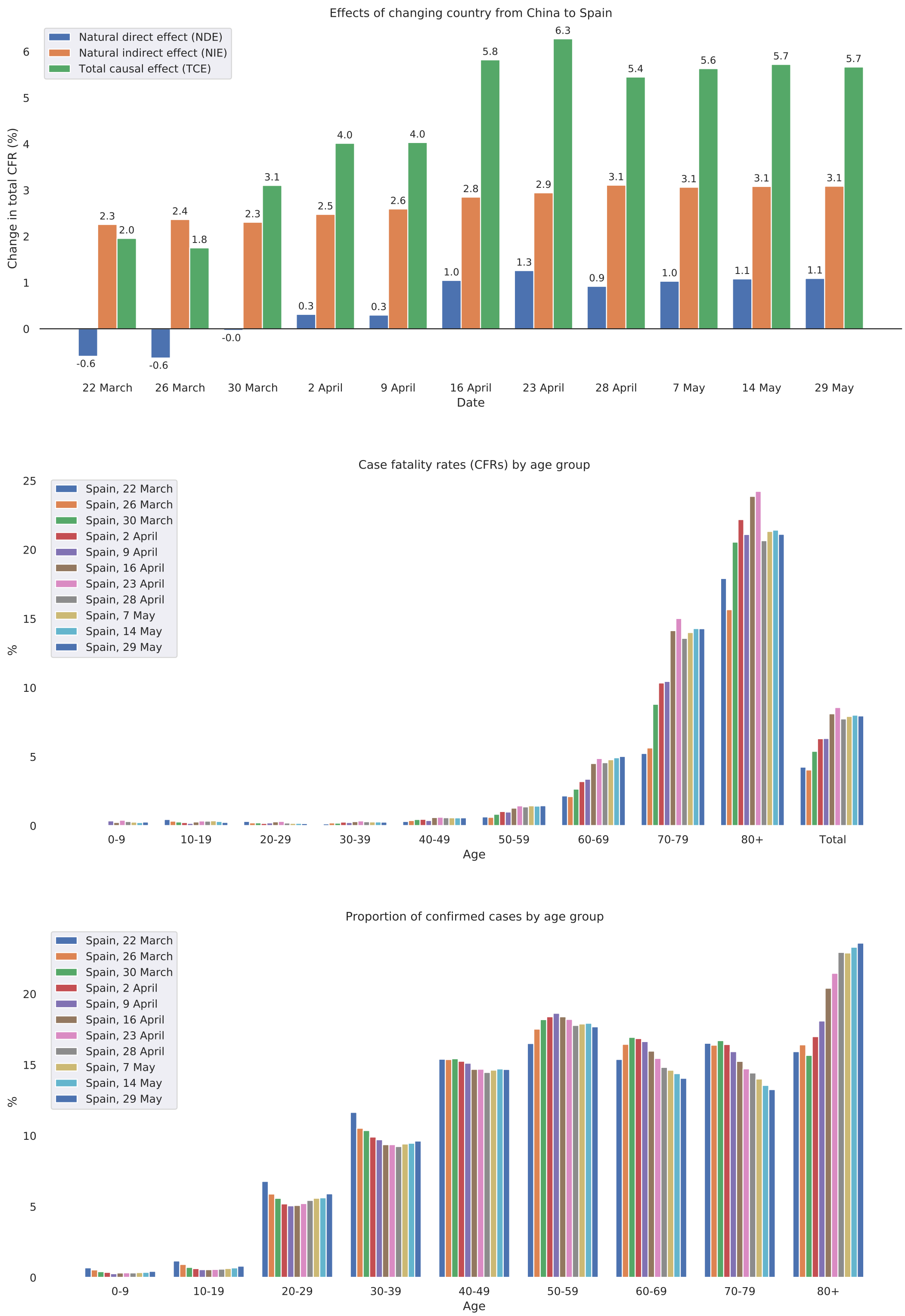


Figure 8: (top) We use different snapshots from Spain to trace TCE, NDE, and NIE of changing country from China to Spain over a time period of 9 weeks, similar to what is shown in Fig. 3 for Italy. We also show the underlying evolution of CFRs by age group (middle) and case demographic (bottom) for the time points considered.

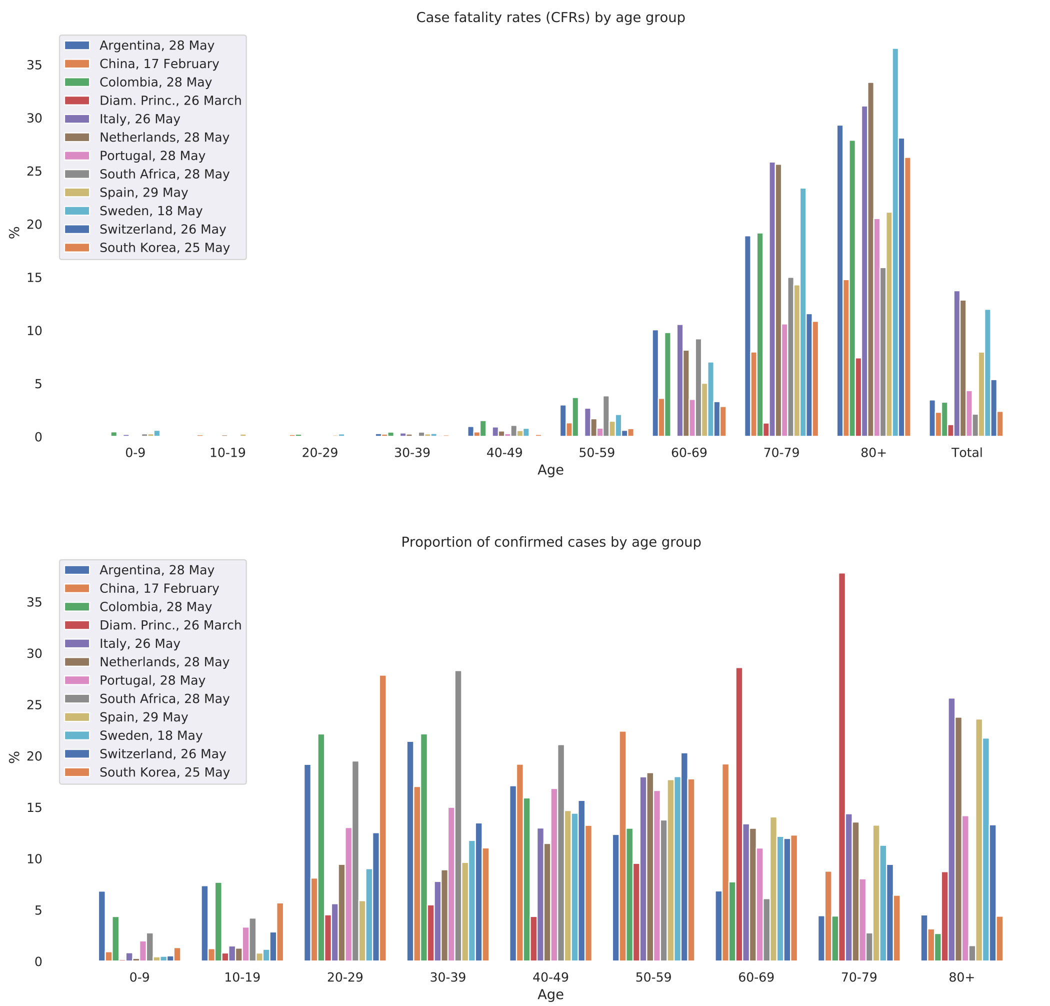


Figure 9: Comparison of CFRs by age group (top) and case demographic (bottom) for all different countries included in our dataset.

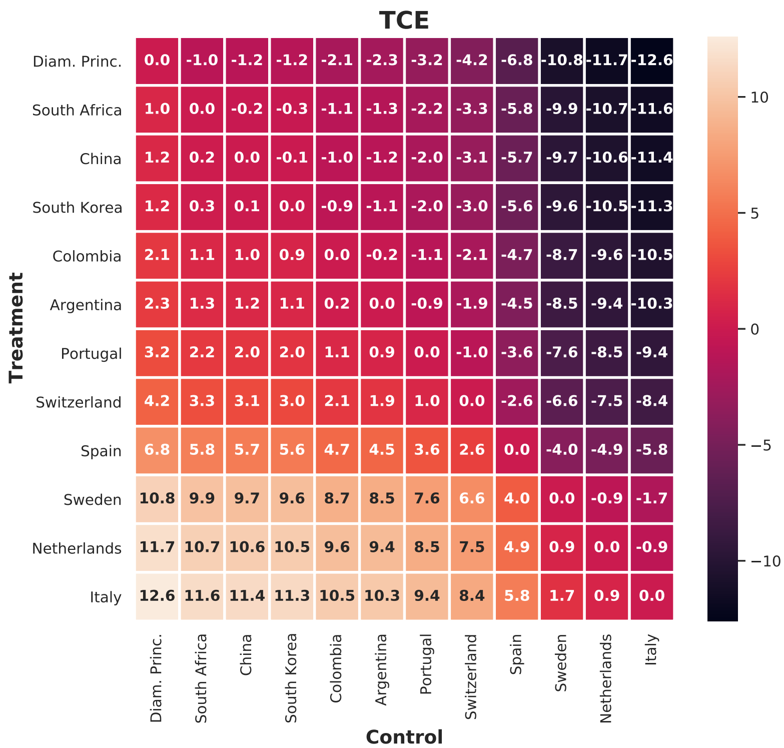


Figure 10: Total causal effects (TCEs) for switching from the control country (columns) to the treatment country (rows). Numbers show the change in total CFR in %, i.e., negative numbers indicate that switching to the treatment country's approach in terms of *both* CFRs by age group *and* case demographic would lead to a decrease in total CFR. Countries are ordered by their average treatment effect over the remaining 11 data points as a control.