NBER WORKING PAPER SERIES

CAUSAL INFERENCE DURING A PANDEMIC: EVIDENCE ON THE EFFECTIVENESS OF NEBULIZED IBUPROFEN AS AN UNPROVEN TREATMENT FOR COVID-19 IN ARGENTINA

Sebastian Calonico
Rafael Di Tella
Juan Cruz Lopez Del Valle

Working Paper 30084
http://www.nber.org/papers/w30084

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2022

We are grateful to Vincent Pons, Amitabh Chandra, Matias Cattaneo, Asim Khwaja, Ariel Stern, Robert Lieli and Pedro Degiovanni, as well as seminar participants at Washington State University, UCSB, UdeSA, Columbia, NYU and Central European University for helpful conversations and comments. For generous help we thank Mariana Butinoff (Centro de Operaciones de Emergencia), Galia Kalayan (Química Luar) and Alexis Doreski (Fundación Respirar). A group of individuals who generously provided help in understanding aspects of health care systems in Argentina and help in locating and gathering data must remain anonymous. Approval for this study was obtained from Columbia's and Harvard's University Institutional Review Boards (IRB-AAAT5728 and IRB20-189, respectively The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w30084.ack

Causal Inference During a Pandemic: Evidence on the Effectiveness of Nebulized Ibuprofen as an Unproven Treatment for COVID-19 in Argentina

Sebastian Calonico, Rafael Di Tella, and Juan Cruz Lopez Del Valle

NBER Working Paper No. 30084
May 2022
JEL No. I1,I18,O3

ABSTRACT

Many medical decisions during the pandemic were made without the support of causal evidence obtained in clinical trials. We study the case of nebulized ibuprofen (NaIHS), a drug that was extensively used on COVID-19 patients in Argentina amidst wild claims about its effectiveness and without regulatory approval. We study data on 5,146 patients hospitalized in 11 health centers spread over 4 provinces, of which a total of 1,019 (19.8%) received the treatment. We find a large, negative and statistically significant correlation between NaIHS treatment and mortality using inverse probability weighting estimators. We consider several threats to identification, including the selection of "low" risks into NaIHS, spillovers affecting patients in the control group, and differences in the quality of care in centers that use NaIHS. While the negative correlation appears to be, broadly, robust, our results are best interpreted as emphasizing the benefits of running a randomized controlled trial and the challenges of incorporating information produced in other, less rigorous circumstances.

Sebastian Calonico
Columbia University
722 West 168th Street
4th Floor
New York, NY 10032
scalonico@gmail.com

Juan Cruz Lopez Del Valle
Universidad de San Andres
Vito Dumas 284
Victoria, Prov B1644BID
Argentina
jclopezdelvalle@gmail.com

Rafael Di Tella Harvard
Business School
Soldiers Field Rd
Boston, MA 02163
and NBER
rditella@hbs.edu

1. Introduction

As COVID-19 cases increased, doctors around the world employed a variety of novel tactics to help their patients. While some of these innovations occurred in the context of clinical trials, others involved a less systematic approach. An interesting case is the treatment of a large number of COVID-19 patients with nebulized ibuprofen (NaIHS) in Argentina. A clinically unsupported treatment, NaIHS was administered to an estimated 61,000 COVID-19 patients amidst wild claims about its effectiveness.[1] Moreover, the diffusion of this treatment took place in spite of explicit regulatory prohibitions, warnings by different professional groups and an initial concern in the international community over the use of non-steroidal anti-inflammatory drugs, which led to an explicit, and much publicized, advice against the use of ibuprofen by the World Health Organization (e.g., see Day, 2020).[2] This gives rise to two important questions: can some of the data and other information generated in such unusual circumstances be recovered and used scientifically? And, more ambitiously, can we overcome at least some of the many challenges to causal inference and make any claim regarding the effectiveness of NaIHS as a treatment against COVID-19?

NaIHS as a treatment for COVID-19 relies on directly delivering a high concentration of ibuprofen in hypertonic saline formulation to the lungs using easily available inhalation devices. Patients, or their legal guardians, provide consent and receive NaIHS three times per day, in addition to standard care. A joint development of a provincial government agency (*CEPROCOR*) and a small pharmaceutical company (*Química Luar*) in the province of Córdoba, the treatment had originally been designed (albeit, not approved) to treat cystic fibrosis but researchers soon conjectured that it might be useful as a treatment for COVID-19 patients (see García et al., 2020).[3]

The Argentine federal regulatory agency (*Administración Nacional de Medicamentos, Alimentos y Teconología Médica, ANMAT*) did not approve the use of NaIHS but on April 2, 2020, the government of the province of Córdoba authorized its use under a novel legal category: "extended compassionate use". This was followed by specific warnings against its use issued by *ANMAT* and by two important professional groups.[4] In spite of this, other provinces soon issued similar authorizations, and

---

[1] As early as May 7, 2020, a leading newspaper in Argentina reported that 5 patients had "successfully" been treated with NaIHS, including two 75 years-old who were seriously ill and needed a respirator. See "Coronavirus en Argentina: investigadores cordobeses prueban con éxito un tratamiento con ibuprofeno" (2020). For reference, to date there have been 128,000 deaths due to COVID-19 in Argentina.

[2] Eventually, the World Health Organization (WHO) withdrew its reservations: "after several initial studies, WHO, the European Medicines Agency (EMA), and the US Food and Drug Administration (FDA) did not advocate against ibuprofen use for COVID-19, but they continue to recommend careful monitoring given the theoretical risk". (Kragholm et al., 2021). See also Drake et al. (2021), which is a large study (n=72,179) demonstrating that the use of non-steroidal anti-inflammatory drugs in COVID-19 patients is not associated with an increased risk of poorer outcomes.

[3] See also "Inhaled Ibuprofen to Treat COVID-19 (Córdoba Trial)" at clinicaltrials.gov, under identification number NCT04382768. Apparently, researchers had shown that high concentrations of salt enhanced the impact of ibuprofen in reducing the infectivity of cystic fibrosis pathogens (see Muñoz et al., 2018). See also Clemente et al., (2021).

[4] See communications on NaIHS by *ANMAT*, the Argentine society of infectiology (*Sociedad Argentina De Infectología, SADI*) and the Argentine society of intensive care (*Sociedad Argentina de Terapia Intensiva, SATI*).

numerous reports of COVID-19 patients treated with NaIHS emerged in local and national media. As late as August 24, 2020, *ANMAT* clarified on its website that no process for a clinical trial had been initiated with the agency and that, lacking a national authorization, transit across jurisdictions (i.e., provinces) was not authorized. Eventually, eight other provinces would issue similar "extended compassionate use" approvals, although NaIHS was often administered in provinces without this legal protection. Soon, a network of approximately 300 pharmacies also produced their own (non-industrial) version of the drug and provided it upon the request of prescribing physicians.[5] NaIHS was eventually administered in all but one of the 24 Argentine provinces (see Calonico et al., 2022a).

NaIHS also seems unusual in the relative paucity of scientific evidence supporting its use. No clinical study was undertaken during the first wave of the pandemic, even as its use spread throughout the country. A phase II clinical trial was authorized by *ANMAT* on July 30, 2021 (after the second wave had ended) but, to the best of our knowledge, has not concluded.[6] Early discussions and evidence appear in García et al. (2020) and Salva et al. (2021) and the results appear promising.[7] Zurita-Lizza & Doreski (2022) provide a possible explanation of the mechanism under play. Other novel medical approaches embraced in Argentina adhered more closely to scientific practices, including research for the phase III study for the Pfizer vaccine (Polack et al., 2020).[8]

In this study, we analyze retrospective data provided by 11 centers (6 of which offered NaIHS at some point in time) on 5,146 patients hospitalized for COVID-19 in Argentina (1,019 of which received industrial NaIHS in addition to standard care, or 19.8% of the sample). We also collect information on the regulatory status of NaIHS in the four provinces in which these 11 health centers are located. Combining these data offers an exceptional window through which we can study the use of a clinically unproven treatment administered in a context of high uncertainty and limited resources during the COVID-19 pandemic. Our approach first explores the decision-making process by modeling the probability of receiving NaIHS as a function of patients' characteristics and other factors. We then compare mortality rates between NaIHS patients and those receiving standard care, controlling for other non-random factors. We discuss possible confounders and other limitations of our study, highlighting the challenges involved in trying to make causal inferences in this context.

---

[5] One (lower bound) estimate of the number of physicians that provided prescriptions for non-industrial NaIHS is 2,247. See also *El Zonda*, 2020. It is hard to classify NaIHS into traditional drug categories (compounded, repurposed, not-legal). The closest is "off-label" which the FDA reserves for cases were an unapproved use of an approved drug (a) is used for a condition for which it is not approved, (b) is given in a different way or (c) is given in a different dose. The use of Ivermectin for COVID-19 is uncommon in that it meets two of them (a and c). NaIHS is a more extreme case since all 3 conditions are met plus it is not widely available as there is an explicit ruling by the regulatory agency (*ANMAT*) against it. Note that the company claims the changes make it a new, unique drug (see *Química Luar*'s communication, 2020).

[6] It is unclear why this is the case. One explanation we received referred to the reduction in the number of COVID-19 patients following the vaccination campaign in Argentina, although this does not preclude a clinical trial. See Argentina's executive disposition DI-2021-5605-APN-ANMAT#MS.

[7] There is little work on the long-term consequences. Note that even with respect to the effects of (non-ibuprofen) nebulized hypertonic saline on infants with acute bronchiolitis, the quality of evidence is only moderate due to substantial clinical heterogeneity between studies (see Zhang et al., 2018). There is also the practical problem of nebulizing patients without spreading the virus. This seems to have been solved by the use of a "helmet" to contain the air leaving the patient.

[8] For evidence on treatments, see Libster et al. (2021) and Santos et al. (2021).

2. Related literature

A large body of work documents variation in treatment propensity (and outcomes) across hospitals in the US, even after controlling for patient risk (see, for example, Skinner, 2011). An interesting case is documented in Chandra & Staiger (2020). They use data on patients that suffered a heart attack and find that hospitals overuse the main treatment (reperfusion therapy) to the point that a group of patients are in fact harmed by the treatment. The variation across hospitals in the survival benefit of the therapy that they estimate is similar to the average treatment effect. Finkelstein et al. (2016) study the causes of variation in health care use by Medicare beneficiaries and are able to separate the role of supply and demand factors.

Clinically unsupported treatments are also sometimes used in high income countries. Cutler et al. (2019) showed that they can account for nearly 12 percent of Medicare total expenditures in the US. Their study finds that patient demand is relatively unimportant in explaining regional variation but that, instead, physician beliefs unsupported by clinical evidence explain 36 percent of end-of-life spending, and 17 percent of U.S. health care spending.

Obviously, the presence of a pandemic may change the preference for conducting clinical trials (over compassionate use) and the WHO, in the aftermath of the 2014 Ebola crisis considered the need of a speedier evaluation tool -which they called the Monitored Emergency Use of Unregistered and Investigational Interventions (see Zuckerman et al., 2021). A related criticism is that our response to the pandemic underplayed financial incentives, for example of pharmaceutical companies in promoting vaccines, and that the preference for scientific evidence emerging in clinical trials over evidence gathered by medical practice is misplaced (see, for example, Blaylock, 2022). Some critics argue that, in the US, patients have initiated lawsuits to force hospitals to administer unproven therapies, including ivermectin (Robertson and Houtz, 2022). Finally, it is worth noting that outside of the US, patients often received off-label therapies as compassionate use during the initial phase of the pandemic (including, hydroxychloroquine, azithromycin, lopinavir-ritonavir, remdesivir and convalescent plasma), but without a control group no firm conclusions were reached (see Kalil, 2020). This was also true evaluating therapies used during the Ebola outbreak. For a recent discussion of the role of real-world data on the evaluation of new drugs, see Concato & Corrigan-Curay (2022).

A remaining puzzle in the episode we study is the time it took to initiate a clinical trial. Budish et al. (2015) study why firms might underinvest in long run research, particularly in health technologies. On the connection between private patenting and public R&D investment in the health sector, see Azoulay et al. (2019). The diffusion of new medical technologies following regulatory approval in the US is discussed in Chandra et al. (2014) (see also Berger et al., 2021).

3. Data

For this study, we obtained retrospective data for 6,262 clinically suspected and PCR positive COVID-19 patients admitted to eleven health centers in Argentina from March 2020 to September 2021. They

were followed until they were discharged or deceased.[9] Data was collected between August 2020 and January 2022 from centers' electronic records or directly from patients' medical records. Sample size was determined by the study time window, no requirements were imposed.

Data on demographics, comorbidities, symptoms, oximetry, treatments and final patient status were gathered. Demographic data include gender and age. Comorbidities were classified into smoking, diabetes, obesity, hypertension, other cardiovascular diseases, pulmonary diseases, chronic kidney disease and cancer, following Center for Disease Control and Prevention guidelines[10]. Oximetry involved respiratory rate and oxygen saturation at admission. Patients received standard care plus at least one of three available treatments: nebulized ibuprofen, dexamethasone and convalescent plasma. The outcome is status at end of stay (discharged or deceased). Complete records on all the dimensions used in this study were available for 5,227 (83.5%) of them, and those with length of stay longer than 56 days (4 times the 14 days disease cycle; Singanayagam et al., 2020) were excluded for a total sample of 5,146 (82.2% of the original).

We constructed two additional sets of relevant variables. First, we defined patients' risk level following Huespe et al. (2022). Second, we defined three measures of center quality. The first is simply centers' use of plasma: the proportion of high or medium risk patients that are treated with plasma in each center. It captures center quality because, at the time, it was believed to be useful against COVID-19 and it was in short supply.[11] The second measure of center quality is a variable called "adherence to the rules," and reflects the extent to which there are visible attempts to conform to best practices. It takes the lowest value (equal to 1) when the center is a public hospital and is administrating NaIHS even when the center is in a province where the government has not issued an extended compassionate use authorization; next (level 2) is when the center is administering NaIHS even when the provincial government has not issued an extended compassionate use authorization and it is a private clinic; next (level 3) is when the center is in compliance with the rules (it is administrating NaIHS in a province where it is authorized under the extended compassionate use label or it is not administering NaIHS; and lastly (level 4) is when both level 3 is obtained and the center is also certified by a well-known quality certification institute (*Instituto Técnico para la Acreditación de Establecimientos de Salud, ITAES*). The third measure of center quality is simply the centers' average of consumers' reports on Google. The three measures are highly correlated (they are reported in appendix's Table A1).

[9] Given the observational nature of the study, ethical approval did not require the patients' written consent.

[10] See "Science Brief: Evidence Used to Update the List of Underlying Medical Conditions Associated with Higher Risk for Severe COVID-19" (2020).

[11] Hospitals that use NaIHS have lower use of plasma (this conclusion holds even if we exclude the hospitals with highest and lowest use of plasma). During this period, a clinical trial was underway and the government sponsored a public campaign to have recovered patients donate plasma. Thus, knowledge of its potential use was not restricted to those working in high quality centers (although, it perhaps required some experience with its use). But it was in short supply, particularly "high-titer" plasma (i.e., plasma with a large concentration of antibodies). Eventually, it was found to be ineffective unless used in the very early stages (Libster et al., 2021). WHO recommended against its use in December 7, 2021 ("WHO recommends against the use of convalescent plasma to treat COVID-19", 2021). Chandra et al. (2014) present evidence of higher quality hospitals adopting a new medical technology earlier. Skinner & Staiger (2015) show how hospitals quickly adopting innovations have better outcomes for their patients. Higher cost hospitals were found to also reach better outcomes by Doyle et al., (2015).

4. Main Results

We now present our main analysis on the role of NaIHS treatment on patients' mortality risk. The lack of a clinical trial makes it difficult to draw strong causal conclusions. Instead, we rely on patients' observable characteristics and other treatments received in order to account for possible selection bias. That is, we use a Selection on Observables assumption, that impose two main conditions:

i. Conditional Independence Assumption (CIA): treatment assignment is as good as random after conditioning on observable characteristics of the participants.
ii. Overlap: individuals with similar observable characteristics are observed in both the treatment and control groups.

While CIA is not directly testable, we will conduct standard validation tests and discuss the sensibility of our findings to potential deviations from it. An additional assumption, not specific to observational studies, is the Stable Unit Treatment Value Assumption (SUTVA). This assumption rules out interference between units, that is, the realized outcome for each participant depends only on the value of the treatment of that unit and not on the treatment or on outcome values of other units (Rubin 2005).

We estimate treatment effects using Inverse Probability Weighting (IPW) methods, where the probability of receiving treatment (the propensity score) is used to reweight the sample such that treated observations with low probability of being treated and untreated observations with high probability of being treated receive larger weights, while treated observations with high probability of being treated and untreated observations with low probability of being treated are underweight. The goal is to make the treatment and control groups comparable based on their observable characteristics. For our main empirical analysis, we use a Doubly Robust IPW (DRIPW) approach that combines IPW with a model for the outcome variable. The appeal of Doubly Robust IPW estimators comes from their efficiency properties.[12] DRIPW estimators are also doubly robust, meaning that only one of the two models must be specified correctly to consistently estimate the treatment effect.

Means of baseline characteristics, their difference and its statistical significance are presented for all patients and by NaIHS treatment status in Table 1. There is a statistically significant imbalance in 12 out of 14 baseline characteristics. Patients with NaIHS treatment were younger on average, suffered less from chronic kidney and pulmonary diseases. They were also more likely to be men, have a higher respiratory rate and lower oxygen saturation, all three baseline characteristics associated with a higher mortality rate in this same sample (Calonico et al., 2022b). Overall, the mortality rate is 18.1%, with 15.5% for patients with NaIHS and 18.8% for patients without NaIHS, a difference of 3.3 percentage points that is statistically significant. Figure 1 splits mortality across centers, without and with NaIHS. In Table 1 we also present *weighted* means differences to check for covariates balance. When weighted

[12] Variations of DRIPW are discussed in Van der Laan & Robins (2003), Bang & Robins (2005), Cattaneo (2010), Farrell (2015), Chernozhukov et al. (2016), Sloczynski & Wooldridge (2017) and Abadie & Cattaneo (2018). Busso et al. (2014) study the finite sample performance of several variants of matching and IPW estimators.

by the inverse of the propensity score, the difference in only 4 (out of the 14 baseline characteristics) is now statistically significant.

Table 2 presents our main results. To begin with, we restrict the sample to centers that use NaIHS. The first column reports the simple, unconditional correlation between mortality and NaIHS. We find a large and statistically significant negative difference of 17 percentage points. In the second column we present the DRIPW estimate of the ATE that combines reweighting (to control for differences in observable characteristics) and a model for the outcome variable. Although smaller in magnitude, we still find a negative and statistically significant correlation (8.7 pp reduction). Figure 2 presents Adjusted Hazard Ratios for the probability of receiving NaIHS in this sample, when restricting to centers that use NaIHS. Findings are in line with Table 1: males and younger patients were more likely to receive the treatment. While obesity shows a positive correlation, there does not seem to be a clear pattern with respect to comorbidities overall.

Threats to identification

Though the conditional independence assumption on which DRIPW rests cannot be tested, in Figure 3a we present (standardized) mean differences in covariates between treatment and control groups for this sample of centers with NaIHS. We find that, once we weigh on the propensity score, the covariates are very well balanced between patients without and with NaIHS. We also check for overlap in the distribution of the propensity score in Figure 3b and indeed observe support for the condition that there must be "comparable" patients both in the treatment and control groups.

Another threat to our identification approach may come from the presence of spillovers from the treatment to control patients within centers that use NaIHS. For example, if physicians provided additional care to treated patients and neglected those that did not receive NaIHS, non-treated patients on the same premises might not provide an appropriate control. We first address this concern by limiting the control group to patients that were treated in these hospitals before NaIHS began to be administered. In Column 3 of Table 2 we compute the DRIPW estimator using a sample where patients treated with NaIHS in centers with NaIHS are compared against patients without NaIHS in centers with NaIHS before the centers start to administer it. The negative association still holds but gets smaller in size and is significant only at the 5% level.

A second approach to deal with this issue involves gathering more data and compare mortality of patients with NaIHS treatment against the mortality of patients treated in centers that never use NaIHS. That is, we excluded patients without NaIHS treatment in centers with NaIHS. Columns 4 and 5 of Table 2 repeat columns 1 and 2 regressions, but for this new sample. The negative association between mortality and NaIHS is now close to zero and no longer statistically significant.

A natural concern with this approach is that centers that never use NaIHS might be different in several dimensions. For instance, they differ in their adherence to existing formal rules. Thus, it is possible that they have higher quality of care in general. For example, the mortality rate of centers that eventually use NaIHS (before they start doing so) is higher than the mortality rate of the centers that

do not administer NaIHS during the same period (36.6% vs 12.5%). Figure 4 also shows that the centers' mean adjusted mortality rates have a negative correlation with one measure of center quality.

A first strategy to take quality into account is to include it directly as a control. We use two proxies: mean use of plasma and the first principal component of the "adherence to the rules" variable and consumers' reports. We take the principal component both for practical reasons (the "adherence to the rules" variable has two levels with only one hospital, see appendix's Table A1) and also because quality measures deliver more accurate estimates when combined (see McClellan & Staiger, 2000). In columns 6 and 7 of Table 2 we replicate column's 5 model but now including our center quality proxies. The negative correlation is now again large and statistically significant.

Previous work has studied the bias in risk-adjusted mortality (Austin, 2015; Finkelstein et al., 2017; Hull, 2018). Doyle et al. (2019) use effectively random assignment of patients to ambulances to make comparisons across patients treated at different hospitals. They find that the survival rates of patients assigned to larger amounts of care is only fractionally higher. In particular, they also find that, even after correcting for patient selection, traditional measures of health center quality, including those based on surveys of patient experience, remain predictive of better outcomes. Our results that control for health center quality are similar when our measure based on consumer reports is used on its own and when it is used in combination with one (or two) of the novel metrics we propose to measure quality (adherence to the rules and mean use of plasma).

During this period there are few approved treatments for COVID-19, so center quality can be expected to play a small role in mortality in general. But, if center quality is expected to play some role, then it is reasonable to expect a bigger impact on the mortality of higher risk patients (for example because these centers might be better at treating patients in the presence of comorbidities).[13] Thus, our second strategy to account for differences in performance between centers consists of splitting the sample into three groups that differ in the role that center quality can be expected to have in explaining mortality. Figure 4 shows how the association between center mean adjusted mortality and quality differ across patients' risk levels. In Table 3 we report DRIPW estimators for samples of patients that differ in their risk level. For both low and medium risk patients there is a negative correlation between mortality and NaIHS. These relationships turn larger and significant when controlling for center quality. For high-risk patients, the basic correlation is positive and insignificant, but it turns negative and statistically significant at the 10% level when a control for center quality is added. Additionally, in Table 4 the association is negative and statistically significant when excluding young patients (under 45 years old) and young patients without comorbidities (and controlling for center quality).

5. Discussion

[13] Physicians working in centers that administer NaIHS claim that the treatment works particularly well on patients with comorbidities (such as obesity). Unfortunately, the nature of our data does not allow us to evaluate this claim so the question of heterogeneity in treatment effects (as in Abrevaya et al., 2014) is left for future work.

Our analysis of the use of nebulized ibuprofen (NaIHS) as a treatment for COVID-19 provides us with a unique window to study how an unproven treatment was used during the pandemic and constitutes one of the largest multicentric studies of hospitalized patients available for developing countries. A first pass at the data restricts attention to hospitals that at some point used NaIHS, and documents a large, statistically significant, mortality reduction that could be attributed to NaIHS treatment. An obvious concern is that this difference could be due to the selection of lower risk patients into NaIHS treatment. However, the unweighted mean differences of the baseline characteristics are not systematically biased towards lower risk patients. More formally, the doubly robust inverse probability weighting estimator confirmed the correlation after correcting for the most plausible sources of potential selection bias.

Another worry is that this basic estimate relies on using patients that were in the same centers but that were not treated with NaIHS as a control group. It is possible that this group was affected by spillovers of the treatment, for example, if physicians were distracted with the treated patients and neglected those that were not treated with NaIHS. To address this issue, we first excluded patients in the control group from the moment that NaIHS begins to be used. In other words, treated patients are compared to patients admitted to these same centers before the arrival of NaIHS. The effect is somewhat smaller and more imprecisely estimated, but remains statistically significant. A more data-intensive procedure excludes all patients not treated with NaIHS in these centers, regardless of the time in which they were admitted, and uses as a control group only patients that were in centers that never use NaIHS. The association between mortality and NaIHS was not different from zero at conventional statistical significance. This result suggests that spillovers within centers using NaIHS could be present, and that the mortality reduction initially attributed to NaIHS could be spurious.

A natural concern is that centers that do not use NaIHS are potentially different from those that use it. During this period, the federal regulator explicitly warned against the use of NaIHS, and some provinces issued unprecedented "extensive compassionate use" authorizations (some centers were not even located within these provinces). Such disregard for formal regulations makes it possible, perhaps even likely, that they do not follow other rules regarding medical procedures. Moreover, the risk-adjusted mortality rate in these centers is lower than that observed amongst non-treated patients in centers that use NaIHS (15.4% versus 29.7%). Thus, the matches provided by the control group might have an artificially low mortality rate if there is either a high quality of care in centers that do not use NaIHS or if these centers select for treatment particularly low risks that are unobservable.

We tackled this through two different strategies. First, measures of center quality were directly included. Once the first measure (centers' use of plasma) is included, there is again a negative and statistically significant correlation between mortality rates and NaIHS. The estimate appears sizeable: 7.6 percentage points relative to a control group mortality rate of 15.6% (for a 48.7% reduction in mortality). This is larger than the previously estimated reduction in mortality obtained when the control group came only from centers that use NaIHS. A second indicator of quality combines a variable that captures "adherence to the rules" with consumers' reports. The estimate is similar to the one obtained in the original sample.

The second strategy exploits a key feature of the pandemic: there were very few treatments available so the quality of care could be expected to play a small role in determining outcomes. Specifically, quality can be expected to have only a relatively small impact on COVID-19 mortality rates, particularly in cases that are not complicated by the existence of comorbidities. In contrast, a relatively higher association between mortality rates and center quality is expected for high-risk patients. Figure 4 shows that this is indeed the case in our sample. When we repeat our main estimate for the three samples (split following risk levels), the correlation between mortality and NaIHS is negative, large and statistically significant (when controlling for center quality) in both the low and medium risk samples. The estimate in the high-risk sample is significant only at the 10% level. Under the assumption that the quality of care is less important for low and medium risk patients, these results are again consistent with the treatment being effective against COVID-19. These estimates should be interpreted with caution as they involve splitting the sample into relatively small groups.

Some informal reports (including from some physicians we interviewed) suggested that better centers often attract relatively richer (or more powerful) patients, who are able to get medical attention even when their symptoms are mild. This points out to the possibility that the low-risk patients from the lower quality centers (that use NaIHS) are different from low-risk patients in centers that do not use NaIHS. There is some evidence in this regard, as the two centers classified as highest quality also have a higher proportion of young patients without comorbidities (15.1%) compared to the rest of the centers (8.9%). Similarly, the sample of low-risk patients (used for the regressions reported in columns 1 and 2 of Table 3) comes disproportionately from the top two hospitals in terms of quality (44% of that sample comes from these two hospitals). If young and young patients without comorbidities have been hospitalized for different (unobservable) reasons, it is of some interest to repeat the regressions excluding them (columns 1 to 4 of Table 4). These results show that the coefficients on NaIHS are again negative, but only significant when controls for center quality are included.

6. Conclusion

A large number of COVID-19 patients in Argentina were treated with a clinically unproven drug, Nebulized Ibuprofen, in spite of explicit warnings by the federal regulator and medical organizations against its use. We collected data on 5,146 patients hospitalized in 11 health centers spread over 4 provinces, of which a total of 1,019 (19.8%) received the treatment. There is a negative and statistically significant correlation between NaIHS treatment and mortality. The most obvious threat to identification, namely the selection of "low" risks into NaIHS treatment, appears unlikely. Spillovers from NaIHS treatment affecting patients in the control group within hospitals is possible so we exploit the multicentric nature of our data to construct a control group from centers that never use NaIHS and are hence unlikely to be affected by spillovers. The estimated effect is negative but insignificant, suggesting the "effect" observed within centers that administer NaIHS is potentially spurious. But centers that adhere to regulations and never use unapproved drugs such as NaIHS are potentially of higher quality. It is thus notable that a similar, negative and statistically significant correlation between mortality and NaIHS use reappears when strategies to control for center quality are considered.

Several limitations, beyond those inherent to any observational study, come from data restrictions. There was no precise information on the timing of NaIHS use (i.e., the precise point at which each patient started the treatment). Other aspects missed in our data include qualitative reports and laboratory results that could have been used by physicians at the time of deciding patients' treatment. Benefits (or harms) from NaIHS could also be reflected in other outcomes not observed in our sample, such as the need for a respirator. Finally, the conclusions turn out to depend on center quality, for which we only have a set of noisy measures.

In the end, the effectiveness of nebulized ibuprofen as a treatment for COVID-19 is suggested by many, but not all, of the estimates obtained. Our results highlight the numerous challenges that arise when making causal inferences in such unusual circumstances. It reinforces the need for randomized controlled trials to more rigorously assess treatment effects, as well as the challenges of incorporating information produced in other, less rigorous circumstances.

References

Abadie, A & Cattaneo, MD (2018). Econometric Methods for Program Evaluation. *Annual Review of Economics* 10:1, 465-503.

Abrevaya, J., Hsu, Y. C., & Lieli, R. P. (2015). Estimating conditional average treatment effects. *Journal of Business & Economic Statistics*, *33*(4), 485-505.

Austin, J. M., Jha, A. K., Romano, P. S., Singer, S. J., Vogus, T. J., Wachter, R. M., & Pronovost, P. J. (2015). National hospital ratings systems share few common scores and may generate confusion instead of clarity. *Health Affairs*, *34*(3), 423-430.

Azoulay, P., Graff Zivin, J. S., Li, D., & Sampat, B. N. (2019). Public R&D investments and private-sector patenting: evidence from NIH funding rules. *The Review of economic studies*, *86*(1), 117-152.

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4), 962-973.

Berger, B., Chandra, A., & Garthwaite, C. (2021). *Regulatory Approval and Expanded Market Size* (No. w28889). National Bureau of Economic Research.

Blaylock R. L. (2022). COVID UPDATE: What is the truth?. *Surgical neurology international*, *13*, 167. https://doi.org/10.25259/SNI_150_2022

Budish, E., Roin, B. N., & Williams, H. (2015). Do firms underinvest in long-term research? Evidence from cancer clinical trials. *American Economic Review*, *105*(7), 2044-85.

Busso, M., DiNardo, J., & McCrary, J. (2014). New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics*, *96*(5), 885-897.

Calonico, S., Di Tella, R. & Lopez Del Valle, J.C. (2022a). The Political Economy of a Miracle Cure: Learning about a Clinically Unproven Treatment for COVID-19 in Argentina. *Mimeo.*

Calonico, S., Di Tella, R. & Lopez Del Valle, J.C. (2022b). Risk factors for in-hospital mortality in COVID-19 patients in Argentina: A competing risk survival analysis. *Mimeo.*

Cattaneo MD (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155(2):138–54

Chandra, A., Malenka, D., & Skinner, J. (2014). The diffusion of new medical technology. *Discoveries in the Economics of Aging*, *389*.

Chandra, Amitabh and Douglas Staiger (2020) "Identifying sources of inefficiency in healthcare", *Quarterly Journal of Economics*, vol. 135, 2, May, Pages 785–843, https://doi.org/10.1093/qje/qjz040.

Chernozhukov, V., Escanciano, J. C., Ichimura, H., Newey, W. K., & Robins, J. M. (2016). Locally robust semiparametric estimation. *arXiv preprint arXiv:1608.00033*.

Clemente CM, Freiberger MI, Ravetti S, Beltramo DM, Garro AG. (2021) An *in silico* analysis of Ibuprofen enantiomers in high concentrations of sodium chloride with SARS-CoV-2 main protease. J Biomol Struct Dyn. Jan 18:1-12. DOI: 10.1080/07391102.2021.1872420. Epub ahead of print. PMID: 33459192; PMCID: PMC7832455.

Concato, J., & Corrigan-Curay, J. (2022). Real-World Evidence-Where Are We Now?. *The New England journal of medicine*.

Coronavirus en Argentina: investigadores cordobeses prueban con éxito un tratamiento con ibuprofeno. (2020, May 7). *Clarín*. Retrieved from https://www.clarin.com/sociedad/coronavirus-argentina-investigadores-cordobeses-prueban-exito-tratamiento-ibuprofeno_05m48GYuWG.html

Cutler, David, Jonathan Skinner, Ariel Dora Stern, and David Wennberg (2019). "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending." *American Economic Journal: Economic Policy* 11, no. 1 (February): 192–221.

Day, M. (2020). Covid-19: ibuprofen should not be used for managing symptoms, say doctors and scientists. *BMJ*. ; 368m1086.

Doyle Jr, J. J., Graves, J. A., Gruber, J., & Kleiner, S. A. (2015). Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy*, *123*(1), 170-214.

Doyle, J., Graves, J., & Gruber, J. (2019). Evaluating measures of hospital quality: Evidence from ambulance referral patterns. *Review of Economics and Statistics*, *101*(5), 841-852.

Drake, TM, Fairfield CJ, Pius R. et al. (2021) "Non-steroidal anti-inflammatory drug use and outcomes of COVID-19 in the ISARIC Clinical Characterisation Protocol UK cohort: a matched, prospective cohort study." *Lancet Rheumatol.* (published online May 7.) https://doi.org/10.1016/S2665-9913(21)00104-1.

El Zonda (2020) "Farmacéuticos ofrecen Ibuprofeno Inhalado Gratis". October 27. *Diario El Zonda*. Retrieved from https://www.diarioelzondasj.com.ar/farmaceuticos-ofrecen-ibuprofeno-inhalado-gratis/

Finkelstein, Amy, Gentzkow, Matthew and Heidi Williams (2016) "Sources of Geographic Variation in Health Care: Evidence From Patient Migration", *Quarterly Journal of Economics*, Volume 131, Issue 4, November 2016, Pages 1681-26, https://doi.org/10.1093/qje/qjw023

Finkelstein, A., Gentzkow, M., Hull, P., & Williams, H. (2017). Adjusting risk adjustment–accounting for variation in diagnostic intensity. *The New England journal of medicine*, *376*(7), 608.

García, N.H., Doreski, P.A., Salva, O., Giler, C., Quinodoz, DC, Guzman, L.G., Muñoz, S., Carrillo, M.N., Ambasch, G., Maldonado, M., Fournier, F. Martinez Picco, C.A. (2020) "Ibuprofeno Inhalado para tratar COVID-19", *Revista Argentina de Terapia Intensiva*. A0147, Sup 2.

García, Néstor H., Daniela J. Porta, Roxana V. Alasino, Sonia E. Muñoz, and Dante M. Beltramo (2020) "Ibuprofen, a traditional drug that may impact the course of COVID-19 new effective

formulation in nebulizable solution", Medical Hypotheses, Med Hypotheses. 2020 Nov; 144: 110079. Published online 2020 Jul 7. doi: 10.1016/j.mehy.2020.110079

Huespe, I., Bisso, I. C., Di Stefano, S., Terrasa, S., Gemelli, N. A., & Las Heras, M. (2022). COVID-19 Severity Index: a predictive score for hospitalized patients. Medicina intensiva, 46(2), 98.

Hull, P. (2018). Estimating hospital quality with quasi-experimental data. *Available at SSRN 3118358*.

Kalil AC. (2020) "Treating COVID-19—Off-Label Drug Use, Compassionate Use, and Randomized Clinical Trials During Pandemics". *JAMA*. 2020;323(19):1897–1898. doi:10.1001/jama..4742

Kragholm, Kristian, Christian Torp Pedersen, Emil Fsobol (2021) "Non-steroidal anti-inflammatory drug use in COVID-19," The Lancet Rheumatology, volume 3, issue 7, E465-E466, JULY 01.

Libster, R., Pérez Marc, G., Wappner, D., Coviello, S., Bianchi, A., Braem, V., ... & Polack, F. P. (2021). Early high-titer plasma therapy to prevent severe Covid-19 in older adults. *New England Journal of Medicine, 384(7), 610-618*.

McClellan, M., & Staiger, D. (2000). Comparing hospital quality at for-profit and not-for-profit hospitals. *The changing hospital industry: Comparing not-for-profit and for-profit institutions*, 93-112.

Muñoz AJ, Alasino RV, Garro AG, Heredia V, García NH, Cremonezzi DC, Beltramo DM (2018) "High concentrations of sodium chloride improve microbicidal activity of ibuprofen against common Cystic Fibrosis pathogens". *Pharmaceuticals* (Basel) 11(2)47 doi: 10.3390/ph11020047

Polack, Fernando P, SJ Thomas, N Kitchin, J Absalon et al. (2020) "Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine", *New England Journal of Medicine*, 2020.

Química Luar communicate. (2020). "Luarprofeno® es igual al ibuprofeno en cápsulas o comprimidos que se encuentra disponible en farmacias?". Retrieved from https://www.quimicaluar.com.ar/luarprofeno/

Robertson, Christopher, and Margaret Houtz (2022). "When Desperate Patients Go to Court for Unproven Treatments—The Battle for Hospital Independence." *New England Journal of Medicine* 386.13: 1200-1202.

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469), 322-331.

Salva O, Doreski PA, Giler CS, Quinodoz DC, Guzmán LG, Muñoz SE, Carrillo MN, Porta DJ, Ambasch G, Coscia E, Diaz JLT, Bueno GD, Fandi JO, Maldonado MA, Peña Chiappero LE, Fournier F, Pérez HA, Quiroga MA, Sala Mercado JA, Martínez Picco C, Beltrán MA, Argañarás LA, Ríos NM, Kalayan GI, Beltramo DM, García NH. (2021). "Reversal of SARS-CoV2-Induced Hypoxia by Nebulized Sodium Ibuprofenate in a Compassionate Use Program". *Infect Dis Ther*. https://doi.org/10.1007/s40121-021-00527-2

Santos JLF, Zanardi P, Alo V, Rodriguez M, Magdaleno F, De Langhe V, Dos Santos V, Murialdo G, Villoldo A, Coria M, Quiros D, Milicchio C, Garcia Saiz E. (2021) "Pulmonary Edema in COVID-19 Treated with Furosemide and Negative Fluid Balance (NEGBAL): A Different

and Promising Approach." *Journal of Clinical Medicine*.; 10(23):5599. https://doi.org/10.3390/jcm10235599

Science Brief: Evidence Used to Update the List of Underlying Medical Conditions Associated with Higher Risk for Severe COVID-19. (2020, Jun). *Center for Disease Control and Prevention.* Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/underlying-evidence-table.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fhcp%2Fclinical-care%2Funderlying-evidence-table.html

Singanayagam, A., Patel, M., Charlett, A., Bernal, J. L., Saliba, V., Ellis, J., ... & Gopal, R. (2020). Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Eurosurveillance*, *25*(32), 2001483.

Skinner, Jonathan (2011) "Causes and consequences of regional variations in health care," *Handbook of Health Economics, Elsevier, vol. 2, pp. 45-93.*

Skinner, J., & Staiger, D. (2015). Technology diffusion and productivity growth in health care. *Review of Economics and Statistics*, *97*(5), 951-964.

Słoczyński, T., & Wooldridge, J. M. (2018). A general double robustness result for estimating average treatment effects. *Econometric Theory, 34(1), 112-133.*

Van der Laan, M. J., Laan, M. J., & Robins, J. M. (2003). *Unified methods for censored longitudinal data and causality*. Springer Science & Business Media.

WHO recommends against the use of convalescent plasma to treat COVID-19. (2021, Dec 7). *World Health Organization.* Retrieved from https://www.who.int/news/item/07-12-2021-who-recommends-against-the-use-of-convalescent-plasma-to-treat-covid-19#:~:text=For%20non%2Dsevere%20COVID%2D19,has%20recovered%20from%20COVID%2D19

Zhang L, Gunther CB, Franco OS, Klassen TP. (2018) Impact of hypertonic saline on hospitalization rate in infants with acute bronchiolitis: A meta-analysis. Pediatr Pulmonol. 2018 Aug;53(8):1089-1095. doi: 10.1002/ppul.24066. Epub 2018 Jun 12. PMID: 29893029.

Zuckerman S, Barlavie Y, Niv Y et al. (2021) "Accessing unproven interventions in the COVID-19 pandemic: discussion on the ethics of 'compassionate therapies' in times of catastrophic pandemics", *Journal of Medical Ethics* Published Online First: 13 October. doi: 10.1136/medethics-2020-106783

Zurita-Lizza, C. & Doreski, A. (2022). "Potential reversal of pulmonary vasoplegia by inhaled ibuprofenate in COVID-19 pneumonia," Clinical and Translational Discovery, Volume 2, Issue 1, March

Table 1.

	All Patients	Unweighted			Weighted		
		With NaIHS	Without NaIHS	Difference	With NaIHS	Without NaIHS	Difference
Male	0.592	0.653	0.577	0.076***	0.598	0.593	0.006
	(0.492)	(0.476)	(0.494)	(0.000)	(0.490)	(0.491)	(0.546)
Age	59.139	58.290	59.349	-1.059*	59.143	59.123	0.019
	(17.606)	(14.673)	(18.254)	(0.085)	(15.081)	(18.038)	(0.953)
Smoking	0.051	0.042	0.053	-0.010	0.056	0.051	0.006
	(0.219)	(0.201)	(0.223)	(0.175)	(0.230)	(0.219)	(0.200)
Diabetes	0.200	0.219	0.195	0.024*	0.204	0.200	0.004
	(0.400)	(0.414)	(0.396)	(0.086)	(0.403)	(0.400)	(0.622)
Obesity	0.191	0.250	0.177	0.073***	0.201	0.191	0.010
	(0.393)	(0.433)	(0.382)	(0.000)	(0.401)	(0.393)	(0.198)
Hypertension	0.409	0.463	0.396	0.068***	0.406	0.408	-0.002
	(0.492)	(0.499)	(0.489)	(0.000)	(0.491)	(0.491)	(0.851)
Other Cardiovascular Diseases	0.125	0.094	0.133	-0.039***	0.116	0.125	-0.009
	(0.331)	(0.292)	(0.340)	(0.001)	(0.320)	(0.331)	(0.157)
Pulmonary Diseases	0.111	0.112	0.111	0.001	0.104	0.111	-0.007
	(0.314)	(0.315)	(0.314)	(0.917)	(0.305)	(0.314)	(0.250)
Chronic Kidney Disease	0.035	0.024	0.038	-0.015**	0.029	0.035	-0.007*
	(0.185)	(0.152)	(0.192)	(0.023)	(0.167)	(0.185)	(0.054)
Cancer	0.054	0.042	0.057	-0.015*	0.042	0.054	-0.012***
	(0.226)	(0.201)	(0.231)	(0.066)	(0.200)	(0.225)	(0.004)
Respiratory Rate	20.755	21.075	20.676	0.399**	20.715	20.801	-0.086
	(5.652)	(3.591)	(6.052)	(0.044)	(3.582)	(6.254)	(0.392)
Oxygen Saturation	93.473	92.114	93.808	-1.694***	93.067	93.296	-0.230**
	(5.606)	(5.401)	(5.606)	(0.000)	(4.340)	(7.151)	(0.049)
Dexamethasone	0.756	0.914	0.717	0.196***	0.756	0.756	-0.000
	(0.429)	(0.281)	(0.450)	(0.000)	(0.430)	(0.429)	(0.958)
Plasma	0.166	0.134	0.173	-0.039***	0.212	0.166	0.045***
	(0.372)	(0.341)	(0.379)	(0.003)	(0.409)	(0.373)	(0.000)
Deceased	0.181	0.155	0.188	-0.033**	0.150	0.193	-0.044***
	(0.385)	(0.362)	(0.391)	(0.015)	(0.357)	(0.395)	(0.000)
Observations	5,146	1,019	4,127	5,146	1,019	4,127	5,146

Notes: Standard errors in parentheses. *** p-value < 0.01, ** p-value < 0.05 and * p-value < 0.1. *Unweighted* refers to raw data. *Weighted* refers to inverse probability weights.

Table 2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Deceased	Deceased	Deceased	Deceased	Deceased	Deceased	Deceased
NaIHS	-0.170***	-0.087***	-0.057**	0.002	-0.018	-0.076***	-0.087***
	(0.020)	(0.024)	(0.028)	(0.013)	(0.014)	(0.022)	(0.019)
Control	0.325	0.297	0.259	0.154	0.154	0.156	0.163
N	1,843	1,843	1,342	4,322	4,322	4,322	4,322
N Treated	1,019	1,019	1,019	1,019	1,019	1,019	1,019
Method	OLS	DRIPW	DRIPW	OLS	DRIPW	DRIPW	DRIPW
Controls	No	Yes	Yes	No	Yes	Yes	Yes
Quarter FE	No	Yes	No	Yes	Yes	Yes	Yes
Center FE	No	Yes	Yes	No	No	No	No
Quality	No	No	No	No	No	MPU	FPC
Sample	1	1	2	3	3	3	3

Notes: Robust standard errors in parentheses. *** p-value < 0.01, ** p-value < 0.05 and * p-value < 0.1. Controls include comorbidities, oximetry and other treatments. Sample 1 uses as a control patients with only standard care in centers that use NaIHS. Sample 2 is the same as sample 1 but excludes patients treated with standard care after the centers begin to use NaIHS. Sample 3 uses as a control patients with only standard care in centers that never use NaIHS. MPU is centers' mean plasma use. FPC is the first principal component that results from combining the "adherence to the rules" variable with consumers' reports.

Table 3.

	(1) Deceased	(2) Deceased	(3) Deceased	(4) Deceased	(5) Deceased	(6) Deceased
NaIHS	-0.002	-0.040***	-0.021	-0.069***	0.041	-0.058*
	(0.010)	(0.015)	(0.021)	(0.023)	(0.032)	(0.035)
Control	0.015	0.017	0.152	0.159	0.328	0.359
N	1175	1175	2136	2136	1011	1011
N Treated	173	173	527	527	319	319
Method	DRIPW	DRIPW	DRIPW	DRIPW	DRIPW	DRIPW
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Center FE	No	No	No	No	No	No
Quality	No	FPC	No	FPC	No	FPC
Sample	3	3	3	3	3	3
Risk Level	Low	Low	Medium	Medium	High	High

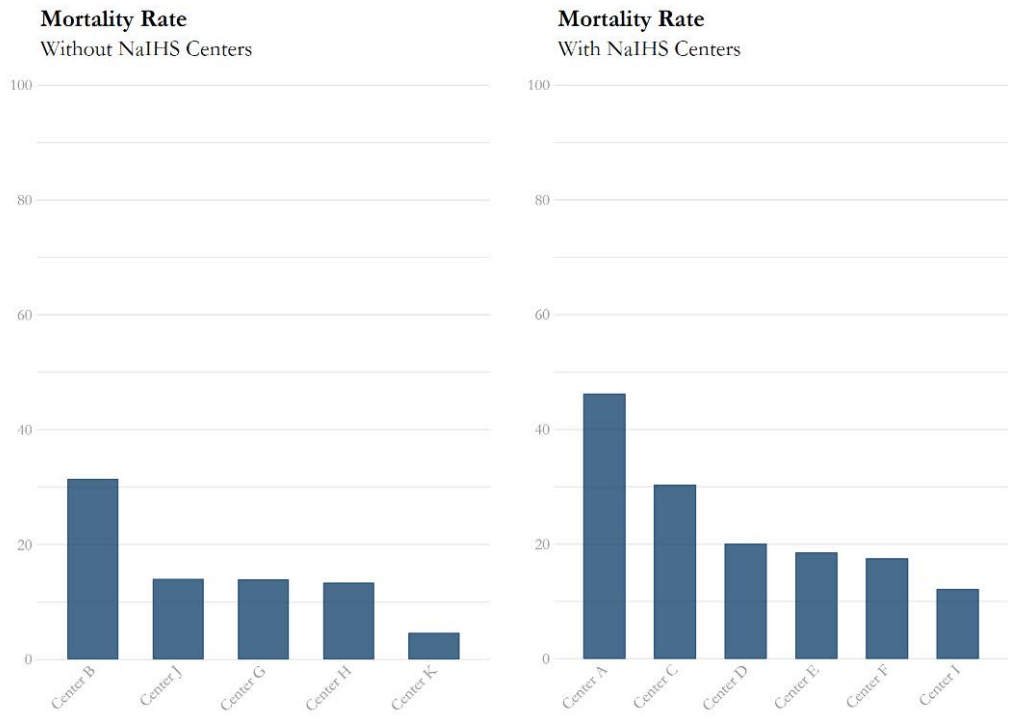
Notes: Robust standard errors in parentheses. *** p-value < 0.01, ** p-value < 0.05 and * p-value < 0.1. Controls include comorbidities, oximetry and other treatments. Sample 3 uses as a control patients with only standard care in centers that never use NaIHS. FPC is the first principal component that results from combining the "adherence to the rules" variable with consumers' reports. Risk Level comes from Huespe et al., (2022). Patients with a score of 7 or more fall within the High Risk Level, patients with a score between 3 and 6 in the Medium Risk Level and patients with a score of 2 or less in the Low Risk Level.

Table 4.

	(1) Deceased	(2) Deceased	(3) Deceased	(4) Deceased
NaIHS	-0.009	-0.085***	-0.010	-0.084***
	(0.017)	(0.020)	(0.016)	(0.018)
Control	0.183	0.196	0.172	0.183
N	3433	3433	3825	3825
N Treated	845	845	931	931
Method	DRIPW	DRIPW	DRIPW	DRIPW
Controls	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Center FE	No	No	No	No
Quality	No	FPC	No	FPC
Sample	3	3	3	3
Risk Level	EY	EY	EYWC	EYWC

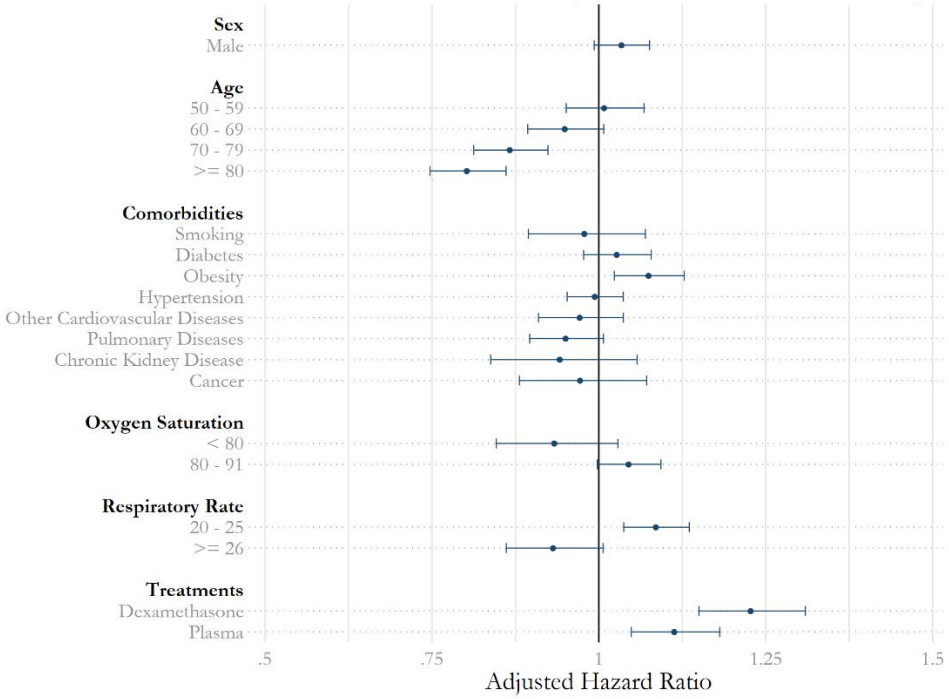
Notes: Robust standard errors in parentheses. *** p-value < 0.01, ** p-value <0.05 and * p-value <0.1. Controls include comorbidities, oximetry and other treatments. Sample 3 uses as a control patients with only standard care in centers that never use NaIHS. FPC is the first principal component that results from combining the "adherence to the rules" variable with consumers' reports. EY stands for Excluding Young (patients that are under 45 years old), while EYWC stands for Excluding Young Without Comorbidities (patients that are under 45 years of age and additionally do not present any comorbidity).

Figure 1.



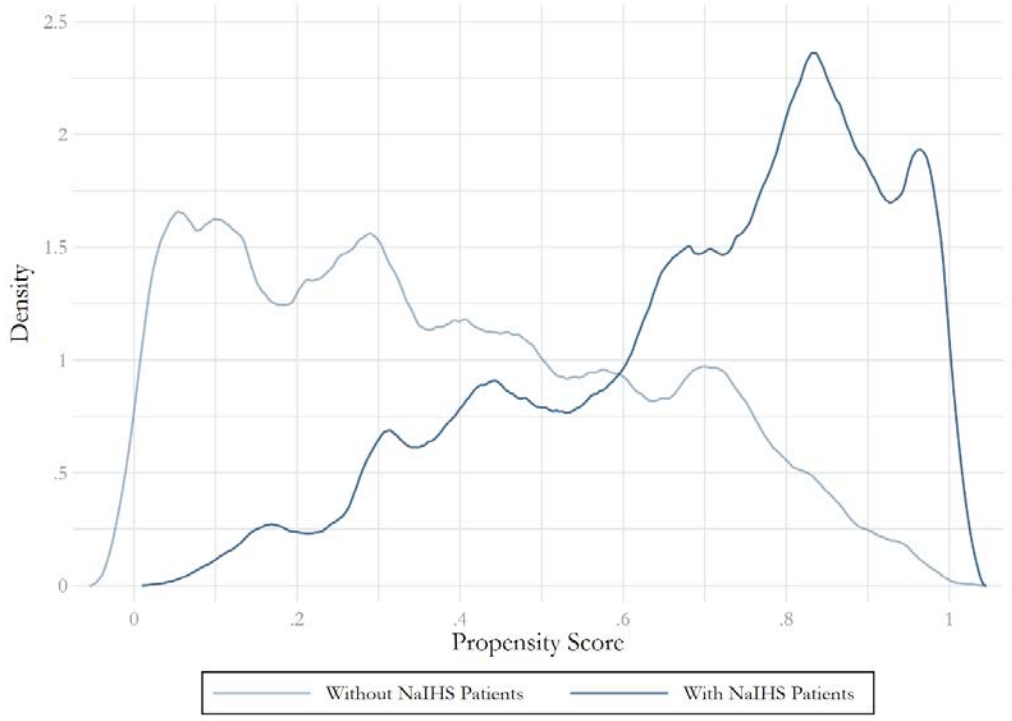
Note: All patients included.

Figure 2.



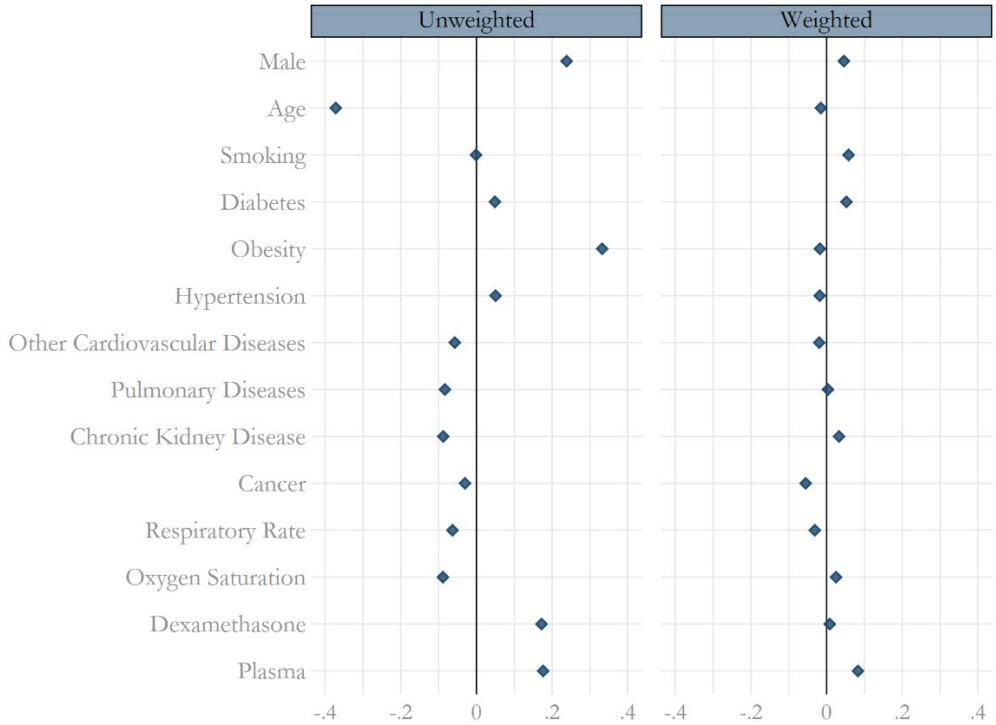
Notes: Adjusted Hazard Ratios come from an OLS regression of NaIHS on covariates, restricting the sample to centers that use NaIHS. 95% confidence intervals. Base category for age is < 50, for oxygen saturation >= 92 and for respiratory rate < 20.

Figure 3a.



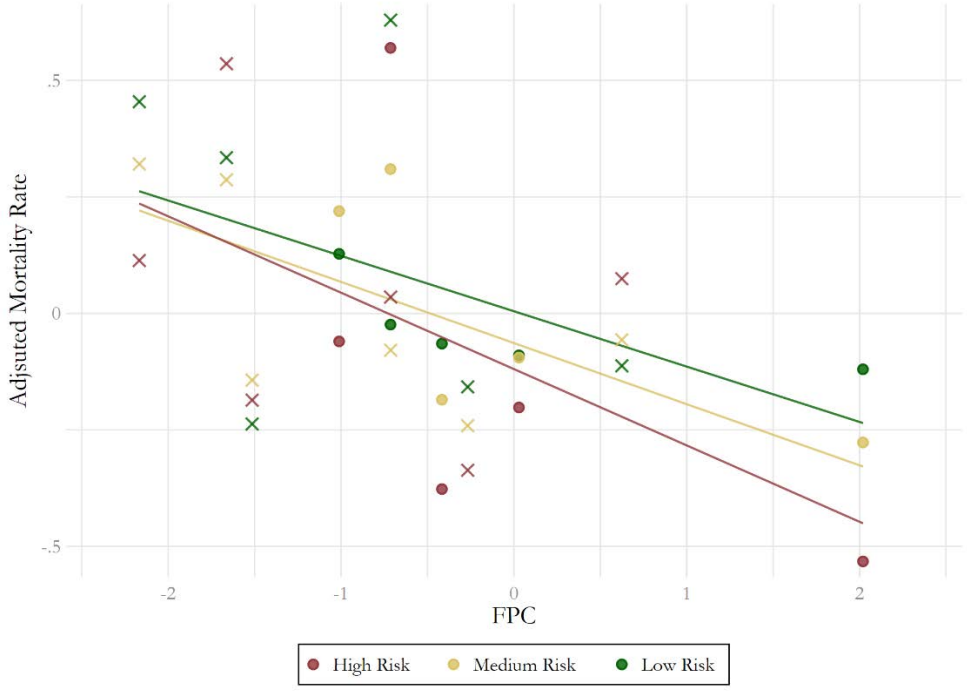
Notes: Sample restricted to centers that use NaIHS.

Figure 3b.



Notes: Standardized differences between patients with and without NaIHS, restricting the sample to centers that use NaIHS. *Unweighted* refers to raw data. *Weighted* refers to inverse probability weights.

Figure 4.



Notes: Adjusted Mortality Rate is centers' mean mortality rate adjusted by covariates. Risk Level comes from Huespe et al. (2022). Patients with a score of 7 or more fall within the High Risk Level, patients with a score between 3 and 6 in the Medium Risk Level and patients with a score of 2 or less in the Low Risk Level. FPC is the first principal component that results from combining the "adherence to the rules" variable with consumers' reports. High Risk has a coefficient of -0.16***, Medium Risk of -0.13** and Low Risk of -0.12**. Dots represent centers that use NaIHS, while crosses represent centers that do not use NaIHS.

Appendix

Table A1.

Quality	Without NaIHS Centers					With NaIHS Centers					
	Center B	Center G	Center H	Center J	Center K	Center A	Center C	Center D	Center E	Center F	Center I
MPU	0.081	0.038	0.708	0.034	0.202	0.101	0.039	0.199	0.099	0.000	0.195
Adh. to the rules	3	3	3	3	4	2	1	3	3	2	3
Consumers' reports	3.4	3.6	3.9	3.2	4.6	3.4	3.7	3.4	4.3	3.5	3.7
FPC	-0.714	-0.417	0.029	-1.011	2.020	-1.664	-2.168	-0.714	0.624	-1.515	-0.268

Note: MPU is centers' mean plasma use. FPC is the first principal component that results from combining the "adherence to the rules" variable with consumer reports.