Investigating the Pathways from World Trade Center Exposures to COVID-19 Severity and Long-Term Sequelae

Bin Yu

May 26 2025

1 Introduction

In late 2019, a new virus called Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) emerged, bringing about a global concern. One of the most important questions to address during the unfolding pandemic is "What are the risk factors for severe illness or death?"

The most frequently mentioned demographic factor increasing the risk for a severe course of disease is higher age (Petrilli et al., 2020; S. J. Rubin et al., 2021), followed by male gender (Jin et al., 2020). Older adults have higher mortality rates and constitute a larger proportion of the patients (Fan et al., 2020; Martín-Sánchez et al., 2020).

Preliminary studies reported severe COVID-19 outcomes in patients with underlying cardiovascular diseases (Chen et al., 2020), arterial hypertension (Yang et al., 2020), and diabetes (Guo et al., 2020). It was also suggested that the presence of underlying respiratory diseases (Aveyard et al., 2021; Sanchez-Ramirez & Mackey, 2020) in general may contribute to severe COVID-19 outcomes. Previous studies have also suggested that the presence of psychiatric conditions such as major depression may increase the risk of COVID-19 mortality in hospitalized individuals (Clouston et al., 2020).

Recently, the Centers for Disease Control and Prevention (CDC) acknowledged that many COVID-19 survivors report symptoms that remain after the infection has disappeared (Bull-Otterson, 2022). Frequent symptoms of PASC include cardiopulmonary problems like shortness of breath and chest pain, along with neuropsychiatric issues such as cognitive fog, fatigue, headaches, and depression (Pandharipande et al., 2023). Studies indicate that six months post-infection, COVID-19 survivors tend to seek medical care more often for respiratory issues, diabetes, and neuropsychiatric conditions than they did before contracting the virus (Richard et al., 2023).

The World Trade Center (WTC) responder cohort, now a middle-aged and aging population, offers an opportunity to trace causal pathways linking pre-existing health conditions with COVID-19 outcomes. Based on the literature, the WTC responder cohort is characterized by high prevalence rates of chronic respiratory conditions such as obstructive airway disease (OAD), chronic obstructive pulmonary disease (COPD), and asthma, as well as significant mental health burdens, including depression (Lhuillier et al., 2022). Studies specifically targeting the WTC cohort have underscored associations between COVID-19 severity and these underlying conditions. For instance, Lhuillier et al. (Lhuillier et al., 2022) found a relationship between COVID-19 severity and conditions such as OAD, depression, and heart disease. Similarly, Waszczuk et al. (Waszczuk et al., 2023) identified asthma polygenic risk scores (PRS) as a factor influencing COVID-19 severity within this group.

The severity of COVID-19 is not only an immediate health concern but also a well-documented predictor of PASC development, as observed in the WTC cohort and broader populations (Babalola et al., 2025). Long-term exposures faced by WTC responders may compound these risks. Severe exposure to environmental hazards at Ground Zero, where responders were subject to acute, high-level toxic exposures, has been linked to respiratory conditions that exacerbate COVID-19 outcomes. Previous studies within this population demonstrate that WTC exposure severity correlates with asthma and other respiratory diseases, including OAD (Alper et al., 2017; Wisnivesky et al., 2011), which are established risk factors for severe COVID-19. However, little is known about the potential effects of acute, extreme toxic exposures—such as those experienced by WTC responders—on COVID-19 disease severity. There is also little evidence about the casual pathway of the development of PASC. Therefore, understanding how unique environmental and occupational exposures among WTC responders may contribute to COVID-19 severity and the development of PASC is critical for informing clinical management, implementing preventive strategies, and addressing health disparities by targeting vulnerable populations affected by occupational hazards.

1.1 Research Hypotheses

Building on the identified research gaps, this study proposes two primary hypotheses corresponding to our two main analyses.

Hypothesis 1: WTC responders who experienced high levels of acute environmental hazard exposures during and after the 9/11 attacks will exhibit more severe COVID-19 outcomes, and this relationship will be mediated by chronic respiratory conditions (e.g., OAD).

Hypothesis 2: Among WTC responders, higher levels of acute toxic exposure will be associated with an increased risk of Long COVID. We further hypothesize that this relationship is mediated in sequence by the development of chronic respiratory conditions (e.g., obstructive airway disease) and by the severity of the initial COVID-19 illness.

2 Data and Methods

This study utilized data from the World Trade Center (WTC) Health Program cohort, which includes 2851 individuals who responded to the 9/11 attacks. Data are collected routinely as part of the WTCHP research efforts, and the analysis of de-identified data was approved by the Institutional Review Boards (IRBs) at SUNY Stony Brook and the University of

Chicago.

2.1 Measures

WTC Exposure:

We used an aggregate measure of exposure intensity based on the approach described by Wisnivesky et al (Wisnivesky et al., 2011). In our aggregate measure, exposure intensity was defined using self-reported duration of exposure to dust, smoke, and debris. Specifically, responders were classified as having:

- *Very High Exposure*: Worked more than 90 days, were exposed to the dust cloud, and worked on the debris pile.
- *High Exposure*: Were exposed to the dust cloud but either worked less than 90 days or did not work on the pile.
- *Intermediate Exposure*: Were not exposed to the dust cloud, yet either worked between 40 and 90 days or did not work on the pile.
- *Low Exposure*: Worked less than 40 days, were not exposed to the dust cloud, and did not work on the pile.

For analysis, the four-level aggregate exposure measure was dichotomized, with low and intermediate exposures coded as 0 and high and very high exposures coded as 1.

Obstructive Airway Disease:

Following the WTC event, a considerable number of responders were subsequently diagnosed with obstructive airway disease (OAD). In our analyses, we treat the documented diagnosis of OAD as a potential mediator of the relationship between WTC exposure and two outcomes: acute COVID-19 severity and post-acute sequelae of SARS-CoV-2 infection (PASC).

COVID-19 Severity:

Participants were classified into four groups based on their COVID-19 symptoms following the NIH COVID-19 clinical spectrum updated in October 2021. Cases were defined as asymptomatic, mild, moderate, and severe: asymptomatic patients tested positive for SARS-CoV-2 but did not report any COVID-19–related symptoms; mild cases included those who experienced at least one symptom without shortness of breath or difficulty breathing and were primarily managed at home despite any initial healthcare visits; moderate cases were characterized by the presence of shortness of breath and/or a diagnosis of lower respiratory

disease (e.g., pneumonia or bronchitis) during clinical evaluation or imaging while maintaining $SpO_2 \geq 94\%$ on room air; and severe cases comprised patients with $SpO_2 \leq 93\%$, a respiratory rate exceeding 30 breaths per minute, a heart rate over 100 beats per minute, or complications such as acute respiratory distress syndrome, septic shock, cardiac dysfunction, or other severe systemic manifestations, including those requiring hospital admission, supplemental oxygen, ICU care, or whose death was attributed to COVID-19. For analysis, the COVID-19 severity measure was dichotomized into a binary variable, with asymptomatic and mild cases coded as 0 and moderate and severe cases coded as 1.

Post-acute Sequelae of SARS-CoV-2 Infection (PASC):

Residual symptoms were defined as any COVID-19–related manifestations that persisted for at least four weeks after the initial symptom onset (Centers for Disease Control and Prevention, 2021b). These symptoms encompassed respiratory issues (such as dyspnea, chest discomfort, cough, and fatigue), central nervous system symptoms (including a loss or reduction in smell/taste, mental fog, dizziness, and vertigo), as well as musculoskeletal complaints. In the analytic sample, 306 participants (31.1%) reported experiencing at least one residual COVID-19 symptom, and a binary variable was created to distinguish between those with and without any residual symptoms.

Confounders:

Other covariates in the analysis included age, education, employment status, race, type of responder, pre-9/11 health problems, and body mass index (BMI). Age was measured as a continuous variable representing the participant's age at the time of the 9/11 events. Education was recoded from self-reported highest educational attainment into four categories: responses of "Grad/Professional School," "BA/BS Degree," and "Associate Degree" were assigned a value of 0; responses of "Some College," "Professional School," and "Technical School" were assigned a value of 1; "High School Graduate" was coded as 2; and both "Some High School" and "Elementary School" were coded as 3. Employment status was derived from the most recent pre–March 1, 2020 record and recoded such that full-time, part-time, sporadic/seasonal, volunteer, and similar work statuses were designated as 0, whereas statuses indicating disability/medical leave, retirement, or unemployment were coded as 1. Race was dichotomized into two groups: White and Non-white. Type of responder was categorized as Traditional (e.g., police, emergency personnel, firefighters), Non-traditional (e.g., doctors, construction workers, etc.). Pre-9/11 health problems were assessed by asking whether participants had experienced any serious illness or injury in the year before 9/11 and whether they had ever had a condition such as depression, alcohol abuse, anxiety, or another psychological problem for which they received care or seriously considered seeking care. Missing values for employment, education, depression, and illness were treated as a

separate category in the analyses since they are not missing at random.

2.2 Analytical Approach

2.2.1 Estimands

In our analysis, the primary outcome is the binary indicator of COVID severity and long COVID, and all observed pretreatment confounders are summarized in the vector X. The exposure is defined as $A =$ Low v.s High WTC Exposure.

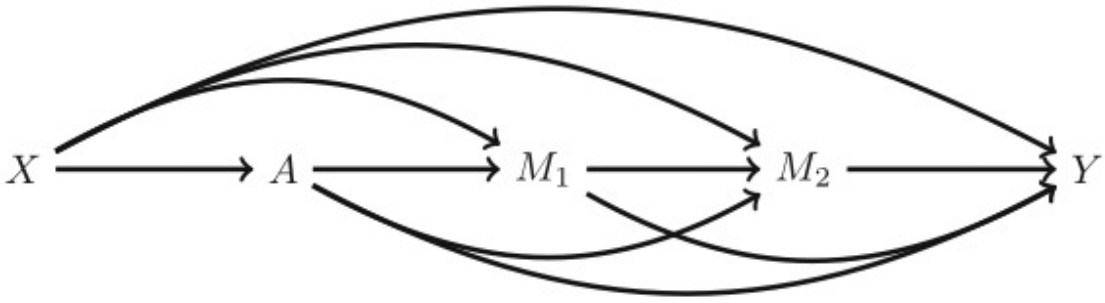

Figure 1: Causal relationships with two causally ordered mediators

Causal mediation analysis was conducted to disentangle the pathways through which the exposure influences the outcome. Our approach leverages the potential outcomes framework (D. B. Rubin, 1974) to provide model-free definitions of these effects (Pearl, 2022). The pathways that transmit the framing effect can be represented by a DAG above. In our analysis, A represents low and high WTC exposure, Y represents long COVID status, M_1 represents represents obstructive airway disease (OAD), and M_2 represents severity of the COVID. In this DAG, several possible causal paths exist from the treatment to the outcome. Now, we define those path-specific effects (PSEs) that we estimate in each of our two mediation models and provide their interpretation in percentage-point terms. The first table covers the OAD to COVID severity model, and the second covers the OAD to COVID severity to long COVID model. Note that in the first table, we use Y to denote M_2 in DAG, which refers to the COVID severity.

Notation	**Definition (Formula)**	**Interpretation**
$\text{PSE}_{A\to Y}(1,0)$	$E[Y(1,\, M_1(0))] \;-\; E[Y(0,\, M_1(0))]$	Expected change (in percentage points) in the probability of moderate/severe COVID if WTC exposure switches from lower level (0) to higher level (1), holding OAD fixed at the level it would naturally take under low exposure.
$\text{PSE}_{A\to M_1\to Y}(1,0)$	$E[Y(1,\, M_1(1))] \;-\; E[Y(1,\, M_1(0))]$	Expected change (in percentage points) in the probability of moderate/severe COVID when WTC exposure is fixed at high level (1) but OAD is switched from its level under low exposure to its level under high exposure, i.e. the portion of the total exposure effect on COVID severity that is mediated solely by OAD.

Table 1: Glossary of Path-Specific Effects (PSEs) for COVID-19 Severity (M_2)in DAG

Notation	Definition (Formula)		Interpretation
$\text{PSE}_{A\to Y}(1,0)$	$E[Y(1,\, M_1(0),\, M_2(0, M_1(0)))]$ $-E[Y(0,\, M_1(0),\, M_2(0, M_1(0)))]$	–	Expected change (in percentage points) in the probability of getting long COIVD if WTC exposure switches from lower level (0) to higher level (1), with both OAD and COVID severity held at the levels they would naturally take under lower level of exposure (0).
$\text{PSE}_{A\to M_2\to Y}(1,0)$	$E[Y(1,\, M_1(0),\, M_2(1, M_1(0)))]$ $-E[Y(1,\, M_1(0),\, M_2(0, M_1(0)))]$	–	Expected change (in percentage points) in the probability of getting long COIVD when WTC exposure is fixed at high (1) but COVID severity is switched from the level it would naturally take under lower level of exposure (0) to the level it would naturally take under high level of exposure (1), while OAD remains at its lower exposure level (0) —i.e. the portion of the total exposure effect on long-COVID risk that is mediated solely by COVID severity.
$\text{PSE}_{A\to M_1\rightsquigarrow Y}(1,0)$	$E[Y(1,\, M_1(1),\, M_2(1, M_1(1)))]$ $-E[Y(1,\, M_1(0),\, M_2(1, M_1(0)))]$	–	Expected change (in percentage points) in the probability of getting long COIVD when WTC exposure is fixed at high level (1) but OAD is switched from the level it would naturally take under lower level of exposure (0) to the level it would naturally take under high level of exposure (1), with acute severity held at its natural level under high level of exposure (1) —i.e. the portion of the total exposure effect on long-COVID risk that is mediated by OAD (including its chain effect on COVID severity).

Table 2: Glossary of Path-Specific Effects (PSEs) for Long COVID

All of the path-specific effects we estimate take the form $\text{PSE} = E\Big[Y(\cdot)\Big] - E\Big[Y(\cdot)\Big]$, where the right-hand side is a difference in expected values of the binary outcome $Y \in \{0, 1\}$. Since $Y \sim \text{Bernoulli}(p) \quad \text{and} \quad E[Y] = P(Y = 1) = p$, each PSE admits a direct probability-scale interpretation: it equals the change in the risk (i.e. the probability) of the event induced by switching the exposure from 0 to 1, while holding mediators at their appropriate "natural" levels. For example, if PSE = 0.05, this corresponds to a 5 percentage-point increase in the probability of the adverse outcome attributable to high versus low exposure.

2.2.2 Identification

The causal effects defined above are identified under sequential ignorability assumptions (Avin et al., 2005; Pearl, 2000; Robins, 2003; Robins & Richardson, 2010), which can be stated as:

$$
\begin{aligned}
&(M_1(a),\, M_2(a, m_1),\, Y(a, m_1, m_2)) \;\perp\!\!\!\perp\; A \;\Big|\; X, \\
&(M_2(a, m_1),\, Y(a, m_1, m_2)) \;\perp\!\!\!\perp\; M_1 \;\Big|\; X,\, A, \\
&Y(a, m_1, m_2) \;\perp\!\!\!\perp\; M_2 \;\Big|\; X,\, A,\, M_1.
\end{aligned}
$$

Here, "$\perp\!\!\!\perp$" denotes statistical independence. The first line asserts that, conditional on baseline covariates X, there is no unmeasured confounding between the joint mediator–outcome process and the exposure A. The second line states that, after adjusting for X and A, there is no unmeasured confounding between the first mediator M_1 and the remaining mediator–outcome pair. The third line requires that, conditional on X, A, and M_1, there is no unmeasured confounding between the second mediator M_2 and the outcome Y.

These assumptions require sufficient overlap (positivity) in the distributions of A, M_1, and M_2 across all strata of X. To mitigate bias, we adjust for an extensive set of pre-exposure covariates X (including demographics, baseline clinical measures, and WTC exposure characteristics) in every model. we also conduct a sensitivity analysis to assess the robustness of our mediation estimates to potential unobserved confounding.

2.2.3 Estimation

To estimate these quantities, we adopted the regression–imputation approach (Zhou & Yamamoto, 2023). Briefly, first, we fit a logistic regression of the binary outcome Y on the exposure A and the full set of pretreatment covariates X. From this baseline model we computed each subject's predicted probability of $Y = 1$ under $A = 0$ and under $A = 1$; these fitted values serve as our initial imputations of the potential outcomes $Y(0)$ and $Y(1)$.

In the second step, we fit a logistic regression of Y on A, X, and both mediators (M_1, M_2). For each individual and each exposure level $d \in \{0, 1\}$, we generated a counterfactual pre-

diction by plugging $A = d$ and the observed mediator values into the fitted model, yielding $\widehat{Y}\big(d,\, M_1(d),\, M_2(d, M_1(d))\big)$. These predictions capture the joint counterfactual risk when both mediators follow their natural values under exposure d.

Next, we regressed the imputed responses from stage 2 on A and X alone. This second-stage regression produces predictions $\widetilde{Y}(d)$ that estimate the marginal counterfactual means $E[Y(d, M_1(d), M_2(d, M_1(d)))]$. We then repeated the same two-stage procedure omitting M_2 from the outcome regression to obtain $\widetilde{Y}(d \mid M_1)$, and finally used the baseline model from stage 1 to represent $\widetilde{Y}(d \mid \text{no mediators})$. Pairwise differences between these three sets of predictions yield the total effect, the natural direct effect (holding both mediators at their natural values under the reference exposure), and the indirect effects through each mediator path.

2.2.4 Sensitivity Analysis

We conducted two complementary sensitivity analyses to assess the robustness of our mediation estimates. In the first analysis, we applied the semi-parametric, multiply-robust estimator of the generalized mediation functional developed by Zhou (Zhou, 2022). Specifically, we implemented both the efficient-influence-function estimating-equation estimator and the targeted maximum likelihood estimator (TMLE) for θ_a. All nuisance functions (the treatment mechanism $P(A \mid X)$, the mediator densities $P(M_k \mid X, A, M_{k-1})$, and the outcome regressions $\mathbb{E}[Y \mid X, A, M_K]$) were estimated via a Super Learner ensemble comprising Lasso regression and random forest, embedded in a 5-fold cross-fitting scheme.

The second sensitivity analysis employs a bias-factor approach (VanderWeele, 2010; Zhou & Yamamoto, 2023) extended to multiple causally ordered mediators. We assume a binary unmeasured confounder U that affects (M_1, M_2) and Y but not A, and impose three simplifying assumptions, including that U is binary, the average "effect" of U on Y, conditional on $(X, A, M_1, \ldots, M_k)$, is constant and denoted γ_k, and the difference in prevalence of U between treated and untreated units, conditional on $(X, M_1, \ldots, M_k)$, is constant and denoted η_k.

Under these assumptions, the biases in the path-specific effects can be written as

$$\text{Bias}\Big[\tau_{A\to Y}\Big] = \gamma_2\,\eta_2, \tag{1}$$

$$\text{Bias}\Big[\tau_{A\to M_1 \rightsquigarrow Y}\Big] = -\,\gamma_1\,\eta_1, \tag{2}$$

$$\text{Bias}\Big[\tau_{A\to M_2\to Y}\Big] = \gamma_1\,\eta_1 \;-\; \gamma_2\,\eta_2. \tag{3}$$

We evaluate robustness by varying (γ_k, η_k) over a grid of plausible values and plotting the resulting bias-adjusted estimates, in particular identifying the combinations that would

attenuate each $\widehat{\tau}$ to zero.

To anchor (γ_k, η_k) in observed data, we exploit an observed binary covariate $Z \in X$ as a proxy for U. First, we fit a linear model of Y on (X, A, M_k, Z) and take the coefficient on Z as an empirical estimate of γ_k. Next, we fit a logistic regression of Z on (X, A, M_k) and use the coefficient on A to estimate η_k. Combining these yields bias-corrected path-specific effects across the sensitivity grid. We find that our substantive conclusions remain stable under moderate departures from the ignorability assumptions.

In the analysis, all continuous variables were standardized prior to analysis by transforming them to Z scores with a mean of zero and standard deviation of one. We used nominal p-value threshold 0.05 and used 1,000 bootstrap replications (percentile method) to derive confidence intervals and p-values. All analyses were conducted in R version 4.1.0.

3 Results

3.1 Descriptive Analysis

Descriptive statistics stratified by the aggregate exposure measure are presented in Table 3 (N = 2225). Among the sample, 1802 responders were classified as having low exposure and 425 as having high exposure. In the low exposure group, 68.0% of responders reported no obstructive airway disease (OAD) while 32.0% reported OAD, compared to 56.0% and 44.0%, respectively, in the high exposure group ($p < 0.001$). With respect to COVID-19 severity, 50.4% of low exposure responders were asymptomatic or experienced mild symptoms, whereas 49.6% had moderate or severe symptoms; in the high exposure group, these proportions were 45.2% and 54.8%, respectively ($p = 0.060$). Furthermore, 27.3% of low exposure responders reported no residual COVID-19 symptoms, while 31.3% of high exposure responders reported at least one residual symptom ($p = 0.112$). The mean age at the time of the 9/11 attacks was 35.43 years (SD = 7.46) for low exposure responders and 36.19 years (SD = 6.69) for high exposure responders ($p = 0.057$). Additional demographic and clinical characteristics—including educational level, employment status, race, type of responder, depression, severe illness, and BMI—are detailed in Table 3. Here, severe illness denotes whether a participant reported having experienced any serious illness or injury before the 9/11 attacks.

Table 3: Descriptive statistics stratified by aggregate exposure measure (N = 2225)

Variable	**Category**	**Low(n = 1802)**	**High(n = 425)**	**p-value**
OAD (%)	No	1225 (68.0)	238 (56.0)	0.001***
	Yes	577 (32.0)	187 (44.0)	
Severity of COVID(%)	Asymptomatic/Mild	908 (50.4)	192 (45.2)	0.060
	Moderate/Severe	894 (49.6)	233 (54.8)	
Long COVID(%)	No	1310 (72.7)	292 (68.7)	0.112
	Yes	492 (27.3)	133 (31.3)	
Age at WTC (mean (SD))	–	35.43 (7.46)	36.19 (6.69)	0.057
Educational Level (%)	Bachelor's	765 (42.5)	173 (40.7)	0.250
	Professional School	532 (29.5)	146 (34.4)	
	High school	250 (13.9)	57 (13.4)	
	Lower than high school	62 (3.4)	9 (2.1)	
	NA	193 (10.7)	40 (9.4)	
Employment Status (%)	Employed	1090 (60.5)	264 (62.1)	0.002**
	Unemployed	443 (24.6)	124 (29.2)	
	NA	269 (14.9)	37 (8.7)	
Race (%)	White	1528 (84.8)	346 (81.4)	0.100
	Non white	274 (15.2)	79 (18.6)	
Type of Responder (%)	Traditional	1275 (70.8)	343 (80.7)	0.001***
	Non-traditional	527 (29.2)	82 (19.3)	
Depression (%)	No	1248 (69.3)	223 (52.5)	0.001***
	Yes	62 (3.4)	13 (3.1)	
	NA	492 (27.3)	189 (44.5)	
Severe Illness (%)	No	1460 (81.0)	344 (80.9)	0.266
	Yes	80 (4.4)	26 (6.1)	
	NA	262 (14.5)	55 (12.9)	
Sex (%)	Female	151 (8.4)	27 (6.4)	0.198
	Male	1651 (91.6)	398 (93.6)	
BMI (mean (SD))	–	30.77 (5.19)	31.29 (5.15)	0.061

Note. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.2 Causal Paths Analysis

Analysis 1: Causal Pathway for COVID Severity

Table 4: Decomposition of Causal Path-Specific Effects for Aggregate WTC Exposure on COVID-19 Severity

Estimator	Effect	Estimate	Std. Err.	95% CI Lower	95% CI Upper	P-value
Pure Imputation Estimator						
	Direct Effect: $A \to Y$	0.035	0.027	−0.017	0.091	0.204
	Indirect Effect: $A \to M_1 \to Y$	0.015	0.006	0.005	0.027	0.004 **
	Total Effect: $A \rightsquigarrow Y$	0.050	0.028	−0.004	0.106	0.068 .
Imputation-Based Weighting Estimator						
	Direct Effect: $A \to Y$	0.035	0.027	−0.017	0.091	0.198
	Indirect Effect: $A \to M_1 \to Y$	0.015	0.006	0.005	0.026	0.004 **
	Total Effect: $A \rightsquigarrow Y$	0.050	0.028	−0.004	0.106	0.068 .
Pure Imputation Estimator with D–M Interaction						
	Direct Effect: $A \to Y$	0.037	0.028	−0.017	0.093	0.178
	Indirect Effect: $A \to M_1 \to Y$	0.013	0.007	0.001	0.029	0.028 *
	Total Effect: $A \rightsquigarrow Y$	0.050	0.028	−0.004	0.106	0.068 .
Imputation-Based Weighting Estimator with D–M Interaction						
	Direct Effect: $A \to Y$	0.037	0.027	−0.020	0.087	0.178
	Indirect Effect: $A \to M_1 \to Y$	0.017	0.006	0.007	0.029	0.034 *
	Total Effect: $A \rightsquigarrow Y$	0.050	0.028	−0.004	0.106	0.068 .

Table 4 shows that, under the pure imputation estimator, moving from low to high WTC exposure yields a direct effect of 0.035 on the probability of moderate/severe COVID-19 (95 % CI [–0.017, 0.091], $p = 0.204$), corresponding to a 3.5 percentage-point increase when OAD is held fixed at its low exposure level. The indirect effect mediated by OAD is 0.015 (95 % CI [0.005, 0.027], $p = 0.004$), indicating a 1.5 percentage-point increase attributable to the OAD pathway when WTC exposure is fixed at high level and OAD is changing from its natural value under low exposure level $M_1(0)$ to high exposure level $M_1(1)$. The total effect is 0.050 (95 % CI [–0.004, 0.106], $p = 0.068$), of which approximately 30 % ($\frac{0.015}{0.050}$) is transmitted through OAD, highlighting a significant mediation channel.

The imputation-based weighting estimator produces virtually identical results: a direct effect of 0.035 (95 % CI [–0.017, 0.091], $p = 0.198$), an indirect effect of 0.015 (95 % CI [0.005, 0.026], $p = 0.004$), and a total effect of 0.050 (95 % CI [–0.004, 0.106], $p = 0.068$). This concordance across methods underscores the robustness of the mediation finding.

Allowing for an exposure–mediator interaction, the pure imputation estimator gives a direct effect of 0.037 (95 % CI [–0.017, 0.093], $p = 0.178$), an indirect effect of 0.013 (95 % CI [0.001, 0.029], $p = 0.028$), and a total effect of 0.050. With imputation-based weighting, the direct effect remains 0.037 (95 % CI [–0.020, 0.087], $p = 0.178$), the indirect effect

increases to 0.017 (95 % CI [0.007, 0.029], $p = 0.034$), and the total effect is again 0.050. Both estimates show that approximately 26-32 % of total effect is transmitted through OAD.

Overall, these analyses consistently demonstrate that OAD mediates a substantive portion of the effect of WTC exposure on COVID-19 severity. The indirect pathways are statistically significant across estimators, highlighting the importance of obstructive airway disease in this causal chain.

Figure 2: Bias-adjusted estimates of the indirect effect with D-M interaction

Figure 2 plots the bias in the estimated indirect effect through M_1 as a function of the sensitivity parameters η_1 (horizontal axis) and γ_1 (vertical axis). The bold contour labeled "0" denotes the bias-adjusted estimates of the indirect effect $A \to M_1 \to Y$ plotted as a function of η_1 and γ_1. The grey area shows the values of η_1 and γ_1 that would reverse the sign of the estimated indirect effects. .

This pattern implies that only an unobserved confounder with either an extreme prevalence difference or an oppositely-signed outcome effect, could reverse the sign of $\tau_{A \to M_1 \to Y}$. In other words, our conclusion of a positive OAD-mediated pathway remains robust to non-extreme unobserved confounding.

Analysis 2: Causal Pathway for Long COVID

Table 5: Decomposition of Causal Path-Specific Effects of Aggregate WTC Exposure on PASC

Estimator	Effect	Estimate	Std. Err.	95% CI Lower	95% CI Upper	P-value
Pure Imputation Estimator without D-M Interaction						
	Direct Effect: $A \to Y$	0.021	0.023	−0.025	0.067	0.376
	Indirect Effect: $A \to M_2 \to Y$	0.008	0.007	−0.005	0.021	0.274
	Indirect Effect: $A \to M_1 \rightsquigarrow Y$	0.008	0.003	0.003	0.015	0.002 **
	Total Effect: $A \rightsquigarrow Y$	0.037	0.025	−0.013	0.086	0.156
Imputation-Based Weighting Estimator without D-M Interaction						
	Direct Effect: $A \to Y$	0.021	0.023	−0.025	0.068	0.374
	Indirect Effect: $A \to M_2 \to Y$	0.008	0.007	−0.006	0.021	0.276
	Indirect Effect: $A \to M_1 \rightsquigarrow Y$	0.008	0.003	0.003	0.015	0.002 **
	Total Effect: $A \rightsquigarrow Y$	0.037	0.025	−0.013	0.086	0.156

Table 6: Type 1 Decomposition of Causal Path-Specific Effects of Aggregate WTC Exposure on PASC

Estimator	Effect	Estimate	Std. Err.	95% CI Lower	95% CI Upper	P-value
Pure Imputation Estimator with D-M Interaction						
	Direct Effect: $A \to Y$	0.028	0.024	−0.021	0.076	0.264
	Indirect Effect: $A \to M_2 \to Y$	0.009	0.007	−0.005	0.023	0.210
	Indirect Effect: $A \to M_1 \rightsquigarrow Y$	0.001	0.005	−0.008	0.012	0.866
	Total Effect: $A \rightsquigarrow Y$	0.037	0.025	−0.013	0.086	0.156
Imputation-Based Weighting Estimator with D-M Interaction						
	Direct Effect: $A \to Y$	0.028	0.024	−0.021	0.076	0.262
	Indirect Effect: $A \to M_2 \to Y$	0.009	0.007	−0.005	0.023	0.216
	Indirect Effect: $A \to M_1 \rightsquigarrow Y$	0.001	0.005	−0.008	0.012	0.862
	Total Effect: $A \rightsquigarrow Y$	0.037	0.025	−0.013	0.086	0.156

Table 7: Type 2 Decomposition of Causal Path-Specific Effects of Aggregate WTC Exposure on PASC

Estimator	Effect	Estimate	Std. Err.	95% CI Lower	95% CI Upper	P-value
Pure Imputation Estimator with D-M Interaction						
	Direct Effect: $A \to Y$	0.017	0.025	−0.033	0.064	0.494
	Indirect Effect: $A \to M_2 \to Y$	0.010	0.007	−0.003	0.025	0.154
	Indirect Effect: $A \to M_1 \rightsquigarrow Y$	0.010	0.004	0.004	0.019	0.001 ***
	Total Effect: $A \rightsquigarrow Y$	0.037	0.025	−0.013	0.086	0.156
Imputation-Based Weighting Estimator with D-M Interaction						
	Direct Effect: $A \to Y$	0.015	0.025	−0.034	0.061	0.544
	Indirect Effect: $A \to M_2 \to Y$	0.011	0.007	−0.003	0.026	0.108
	Indirect Effect: $A \to M_1 \rightsquigarrow Y$	0.011	0.004	0.003	0.020	0.002 **
	Total Effect: $A \rightsquigarrow Y$	0.037	0.025	−0.013	0.086	0.156

Tables 5–7 display the decomposition of causal path-specific effects of aggregate WTC exposure on PASC under three analytic scenarios.

First, when no exposure–mediator (D–M) interaction is included in the model (Table 5), both the pure imputation estimator and the imputation-based weighting estimator yield virtually identical results. The direct effect of exposure on PASC is small and non-significant (0.021; 95% CI [–0.025, 0.067], $p = 0.376$), the indirect effect through severity ($A \to M_2 \to Y$) is also non-significant (0.008; 95% CI [–0.005, 0.021], $p = 0.274$). In contrast, the indirect effect via obstructive airway disease ($A \to M_1 \rightsquigarrow Y$) is positive and statistically significant (0.008; 95% CI [0.003,0.015], $p < 0.01$), leading to a total effect of 0.037 (95% CI [-0.013, 0.086] ,$p = 0.156$).

Therefore, under the assumption of no D–M interaction—i.e. that the effect of exposure of WTC on PASC is constant across mediator levels, we only have statistically significant evidence of an effect mediated by OAD, which represents 0.8 percentage points of probability of PASC when OAD is changing from its natural value under low exposure level $M_1(0)$ to high exposure level $M_1(1)$, while holding WTC exposure fixed at high level and COVID severity M_2 at its natural value under each scenario. Approximately 21% ($\frac{0.008}{0.037}$) of total effect is transmitted through this pathway.

Now we allow for a D–M interaction in the outcome model and compare two alternative decompositions:

Next, allowing for an exposure–mediator interaction in the outcome model and using the Type I decomposition (Table 6), which fixes M_1 and M_2 at their counterfactual values under high exposure, the pure imputation estimator yields a same total effect, but a slightly larger direct effect of 0.028 (95 % CI [–0.021, 0.076]; $p = 0.264$), while the indirect effect via OAD attenuates to 0.001 (95 % CI [–0.008, 0.012]; $p = 0.866$) and the pathway through COVID severity remains non-significant (0.009, 95% CI [–0.005, 0.023], $p = 0.210$)).

Under an exposure–mediator interaction and using the Type II decomposition (Table 7), which fixes both mediators at their low exposure counterfactuals $M_1(0)$ and $M_2(0, M_1(0))$, the pure imputation estimator yields a direct effect of 0.017 (95 % CI [–0.033, 0.064]; $p = 0.494$), an indirect effect via COVID severity of 0.010 (95 % CI [–0.003, 0.025]; $p = 0.154$), and a siginificant indirect effect via OAD of 0.010 (95 % CI [0.004, 0.019]; $p < 0.001$), for a total effect of 0.037 (95 % CI [–0.013, 0.086]; $p = 0.156$). The effect mediated by OAD consists 27% of the total effect. The imputation-based weighting estimator gives similar results.

The divergence between Type I and Type II decompositions arises because, in the presence of an exposure–mediator interaction, the contribution of A to Y depends on the level at which the mediator is held fixed. Type I evaluates the contrast $Y(1, M(0)) - Y(0, M(0))$,

whereas Type II evaluates $Y(1, M(1)) - Y(0, M(1))$, leading to different estimates when A $\times$ M terms are present.

In our scenario, the only statistically significant mediated pathway emerges under the Type II decomposition, the PSE $A \to M_1 \rightsquigarrow Y$ fixes the exposure and mediator M_2 at their low exposure counterfactuals $Y(0)$ and $M_2(0,\ M_1(0))$. Concretely, Type II isolates

$$E\Big[Y\Big(0, M_1(1), M_2\{0, M_1(1)\}\Big)\Big] - E\Big[Y\Big(0, M_1(0), M_2\{0, M_1(0)\}\Big)\Big],$$

which here equals 0.010 (95 % CI [0.004, 0.019]; $p < 0.001$). This reflects the change of percentage points in PASC probability when OAD is changing from its low exposure level to high exposure level, while holding $A = 0$ and severity at its natural value given each OAD status and low exposure level.

By contrast, the Type I decomposition fixes both exposure and mediator M_2 at their higher exposure level counterfactuals $Y(1)$ and $M_2(1, M_1(0))$, and yields an indirect effect mediated by OAD of only 0.001 (95 % CI [–0.008, 0.012]; $p = 0.866$). Here Type I isolates

$$E\Big[Y\Big(1,\ M_1(1),\ M_2\{1,\ M_1(1)\}\Big)\Big] \ -\ E\Big[Y\Big(1,\ M_1(0),\ M_2\{1,\ M_1(0)\}\Big)\Big]$$

it captures the OAD-mediated increase that arises when OAD itself is changing from low to high-exposure level, while both COVID severity and exposure fixed at high exposure counterfactuals.

Thus, when we allow for an exposure–mediator interaction, only the Type II decomposition—where we hold exposure and mediators at their low-exposure counterfactuals $Y(0)$ and $M_2(0,\ M_1(0))$ while switching OAD status from 0 to 1 yields a statistically significant OAD-mediated effect (0.010, 95 % CI [0.004, 0.019]; $p < 0.001$) By contrast, the Type I decomposition—where we fix mediators at their high-exposure levels $Y(1)$, $M_2(1,\ M_1(1))$ returns an OAD-mediated effect of only 0.001 (95 % CI [–0.008, 0.012]; $p = 0.866$). In other words, the indirect pathway via OAD is detectable among individuals with low exposure (or with low mediator levels), that only when we benchmark at those low-exposure mediator values does the indirect effect become detectable.

Table 8: Type I Decomposition of Causal Path-Specific Effects of Aggregate WTC Exposure on PASC (Sensitivity Analysis)

Estimator	Effect	Estimate	95% CI Lower	95% CI Upper
Estimating Equation (EIF)				
	Total Effect	0.0421	−0.0112	0.0954
	via OAD	0.0050	−0.0047	0.0148
	via Severity	0.0126	−0.0015	0.0267
	Direct Effect	0.0245	−0.0239	0.0729
TMLE				
	Total Effect	0.0437	−0.0091	0.0965
	via OAD	0.0031	−0.0064	0.0125
	via Severity	0.0271	0.0127	0.0415
	Direct Effect	0.0135	−0.0344	0.0615

Table 9: Type II Decomposition of Causal Path-Specific Effects of Aggregate WTC Exposure on PASC (Sensitivity Analysis)

Estimator	Effect	Estimate	95% CI Lower	95% CI Upper
Estimating Equation (EIF)				
	Total Effect	0.0421	−0.0112	0.0954
	via OAD	0.0093	0.0030	0.0153
	via Severity	0.0112	−0.0020	0.0245
	Direct Effect	0.0215	−0.0260	0.0690
TMLE				
	Total Effect	0.0437	−0.0091	0.0965
	via OAD	0.0030	−0.0030	0.0090
	via Severity	0.0156	0.0016	0.0296
	Direct Effect	0.0251	−0.0260	0.0760

We conducted complementary sensitivity analyses using the semi-parametric, multiply-robust estimator of the generalized mediation functional of Zhou (Zhou, 2022). Both the efficient-influence-function estimating-equation (EIF) estimator and the targeted maximum likelihood estimator (TMLE) for θ_a were implemented, with all nuisance functions estimated via a Super Learner ensemble (Lasso, random forest) in a 5-fold cross-fitting scheme.

These sensitivity estimates (total effect ≈ 0.04, direct effect ≈ 0.25) are broadly in line with our primary results, although they tend to attenuate the OAD pathway, rendering it non-significant under TMLE in both decompositions and under EIF for the Type I decomposition. However, the EIF in Type II decomposition shows that the significant indirect effect mediated by OAD is significant, which aligns with our main results. Overall, the pattern of a modest total effect and small mediated effect is preserved, demonstrating that our findings

are relatively robust.

Figure 3: Bias–adjusted estimates of the indirect effect $A \to M_1 \to Y$ as a function of the sensitivity parameters η_1 (difference in prevalence of U) and γ_1 (effect of U on the outcome). Contour lines are labeled with the bias value; the bold "0" contour marks where the indirect effect estimate would remain unchanged.

Figure 3 plots the bias in the estimated indirect effect through M_1 as a function of the sensitivity parameters η_1 (horizontal axis) and γ_1 (vertical axis). We could see that, like figure 2, only an unobserved confounder with an extreme prevalence difference or an oppositely-signed outcome effect could reverse the sign of $\tau_{A \to M_1 \rightsquigarrow Y}$. In other words, our conclusion of a positive OAD-mediated pathway remains robust to non-extreme unobserved confounding.

4 Conclusion

In this study of 2,225 WTC responders, we combined aggregate exposure measures with sequential causal mediation analysis to disentangle the pathways linking 9/11 work exposures to COVID-19 severity and long-term post-COVID symptoms (PASC). Descriptive analyses revealed higher rates of OAD and more severe COVID-19 among highly exposed responders. Causal path analysis showed that a change in OAD status mediates a 1.5-percentage-point increase in the probability of moderate/severe COVID-19 (indirect = 0.015; 95 % CI [0.005, 0.027]), corresponding to 30 % of the total 5.0-percentage-point effect of higher WTC exposure. For PASC, OAD mediates a 0.8-percentage-point increase in the probability of long

COVID (indirect = 0.008; 95 % CI [0.003, 0.015]), approximately 22 % of the total 3.7-percentage-point effect of higher exposure. Notably, this mediated effect is significant only among those with lower WTC exposure, suggesting that the impact of exposure on PASC varies by exposure level. In contrast, both the direct effect of WTC exposure and the indirect effect via COVID-19 severity on PASC remain non-significant.

Future research should further elucidate the interplay between environmental exposures, host characteristics, and viral dynamics to inform more effective preventive strategies and public health policies for WTC responders.

Data and Code Availability Statement

This work uses sensitive health data and other information of patients enrolled in the Stony Brook University World Trade Center (SBU-WTC) Health and Wellness Monitoring Program. Therefore, I only submit the deidentifying and clean data. All codes for data cleaning and analysis have been attached.

References

Alper, H. E., Yu, S., Stellman, S. D., & Brackbill, R. M. (2017). Injury, intense dust exposure, and chronic disease among survivors of the World Trade Center terrorist attacks of September 11, 2001. *Injury Epidemiology*, *4*(1), 17. https://doi.org/10.1186/s40621-017-0115-x

Aveyard, P., Gao, M., Lindson, N., Hartmann-Boyce, J., Watkinson, P., Young, D., Coupland, C. A., San Tan, P., Clift, A. K., & Harrison, D. (2021). Association between pre-existing respiratory disease and its treatment, and severe COVID-19: A population cohort study. *The lancet Respiratory medicine*, *9*(8), 909–923.

Avin, C., Shpitser, I., & Pearl, J. (2005). Identifiability of path-specific effects.

Babalola, T. K., Clouston, S. A., Sekendiz, Z., Chowdhury, D., Soriolo, N., Kawuki, J., Meliker, J., Carr, M., Valenti, B. R., & Fontana, A. (2025). SARS-COV-2 re-infection and incidence of post-acute sequelae of COVID-19 (PASC) among essential workers in New York: A retrospective cohort study. *The Lancet Regional Health–Americas*, *42*.

Bull-Otterson, L. (2022). Post–COVID conditions among adult COVID-19 survivors aged 18–64 and 65 years—United States, March 2020–November 2021. *MMWR. Morbidity and Mortality Weekly Report*, *71*.

Chen, R., Liang, W., Jiang, M., Guan, W., Zhan, C., Wang, T., Tang, C., Sang, L., Liu, J., & Ni, Z. (2020). Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in China. *Chest*, *158*(1), 97–105.

Clouston, S. A., Luft, B. J., & Sun, E. (2020). History of premorbid depression is a risk factor for COVID-related mortality: Analysis of 1,387 COVID+ patients [Publisher: Cold Spring Harbor Laboratory Press]. *medRxiv*, 2020–12.

Fan, H., Tang, X., Song, Y., Liu, P., & Chen, Y. (2020). Influence of COVID-19 on cerebrovascular disease and its possible mechanism. *Neuropsychiatric Disease and Treatment*, 1359–1367.

Guo, W., Li, M., Dong, Y., Zhou, H., Zhang, Z., Tian, C., Qin, R., Wang, H., Shen, Y., & Du, K. (2020). Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes/metabolism research and reviews*, *36*(7), e3319.

Jin, J.-M., Bai, P., He, W., Wu, F., Liu, X.-F., Han, D.-M., Liu, S., & Yang, J.-K. (2020). Gender differences in patients with COVID-19: Focus on severity and mortality. *Frontiers in public health*, *8*, 545030.

Lhuillier, E., Yang, Y., Morozova, O., Clouston, S. A. P., Yang, X., Waszczuk, M. A., Carr, M. A., & Luft, B. J. (2022). The Impact of World Trade Center Related Medical

Conditions on the Severity of COVID-19 Disease and Its Long-Term Sequelae. *International Journal of Environmental Research and Public Health*, *19*(12). https://doi.org/10.3390/ijerph19126963

Martín-Sánchez, F. J., Del Toro, E., Cardassay, E., Valls Carbo, A., Cuesta, F., Vigara, M., Gil, P., Lopez Picado, A. L., Martinez Valero, C., & Miranda, J. D. (2020). Clinical presentation and outcome across age categories among patients with COVID-19 admitted to a Spanish Emergency Department. *European geriatric medicine*, *11*, 829–841.

Pandharipande, P., Roberson, S. W., Harrison, F. E., Wilson, J. E., Bastarache, J. A., & Ely, E. W. (2023). Mitigating neurological, cognitive, and psychiatric sequelae of COVID-19-related critical illness. *The Lancet Respiratory Medicine*, *11*(8), 726–738.

Pearl, J. (2000). Causality: Models, reasoning, and inference.

Pearl, J. (2022). Direct and indirect effects. In *Probabilistic and causal inference: The works of Judea Pearl* (pp. 373–392).

Petrilli, C. M., Jones, S. A., Yang, J., Rajagopalan, H., O'Donnell, L., Chernyak, Y., Tobin, K. A., Cerfolio, R. J., Francois, F., & Horwitz, L. I. (2020). Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: Prospective cohort study. *bmj*, *369*.

Richard, S. A., Pollett, S. D., Fries, A. C., Berjohn, C. M., Maves, R. C., Lalani, T., Smith, A. G., Mody, R. M., Ganesan, A., & Colombo, R. E. (2023). Persistent COVID-19 symptoms at 6 months after onset and the role of vaccination before or after SARS-CoV-2 infection. *JAMA network open*, *6*(1), e2251360–e2251360.

Robins, J. M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. *Highly structured stochastic systems*, 70–82.

Robins, J. M., & Richardson, T. S. (2010). Alternative graphical causal models and the identification of direct effects. *Causality and psychopathology: Finding the determinants of disorders and their cures*, *84*, 103–158.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.

Rubin, S. J., Falkson, S. R., Degner, N. R., & Blish, C. (2021). Clinical characteristics associated with COVID-19 severity in California. *Journal of Clinical and Translational Science*, *5*(1), e3.

Sanchez-Ramirez, D. C., & Mackey, D. (2020). Underlying respiratory diseases, specifically COPD, and smoking are associated with severe COVID-19 outcomes: A systematic review and meta-analysis. *Respiratory medicine*, *171*, 106096.

VanderWeele, T. J. (2010). Bias formulas for sensitivity analysis for direct and indirect effects [Publisher: LWW]. *Epidemiology*, *21*(4), 540–551.

Waszczuk, M. A., Morozova, O., Lhuillier, E., Docherty, A. R., Shabalin, A. A., Yang, X., Carr, M. A., Clouston, S. A. P., Kotov, R., & Luft, B. J. (2023). Polygenic risk scores for asthma and allergic disease associate with COVID-19 severity in 9/11 responders. *PLOS ONE*, *18*(3), e0282271. https://doi.org/10.1371/journal.pone.0282271

Wisnivesky, J. P., Teitelbaum, S. L., Todd, A. C., Boffetta, P., Crane, M., Crowley, L., de la Hoz, R. E., Dellenbaugh, C., Harrison, D., Herbert, R., Kim, H., Jeon, Y., Kaplan, J., Katz, C., Levin, S., Luft, B., Markowitz, S., Moline, J. M., Ozbay, F., ... Landrigan, P. J. (2011). Persistence of multiple illnesses in World Trade Center rescue and recovery workers: A cohort study. *The Lancet*, *378*(9794), 888–897. https://doi.org/10.1016/S0140-6736(11)61180-X

Yang, J., Zheng, Y., Gou, X., Pu, K., Chen, Z., Guo, Q., Ji, R., Wang, H., Wang, Y., & Zhou, Y. (2020). Prevalence of comorbidities in the novel Wuhan coronavirus (COVID-19) infection: A systematic review and meta-analysis. *Int J Infect Dis*, *94*(1), 91–95.

Zhou, X. (2022). Semiparametric estimation for causal mediation analysis with multiple causally ordered mediators. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *84*(3), 794–821.

Zhou, X., & Yamamoto, T. (2023). Tracing causal paths from experimental and observational data. *The Journal of Politics*, *85*(1), 250–265.