Original Research Article

DIGITAL HEALTH

DIGITAL HEALTH
Volume 10: 1–10

DOI: 10.1177/20552076241261892
journals.sagepub.com/home/dhj

Sage

Change in disease burden associated with influenza and air pollutants during the COVID-19 pandemic in Hong Kong

Yanwen Liu[1,^], Xie Jingyu[2,^], Cai Cihan[1,^], Hilda Tsang[3], Shuya Lu[3], Daihai He[1] and Lin Yang[3]

Abstract

Objectives: This study aimed to estimate the variation in disease burden associated with air pollutants and other respiratory viruses during the COVID-19 pandemic.

Methods: We adopted a machine learning approach to calculate the excess mortality attributable to air pollutants and influenza, during the pre-pandemic and pandemic period.

Results: In the first 2 years of the COVID-19 pandemic, there were 8762 (95% confidence interval, 7503–9993), and 12,496 (11,718–13,332) excess all-cause deaths in Hong Kong. These figures correspond to 117.4 and 167.9 per 100,000 population, and 12.6% and 8.5% of total deaths in 2020 and 2021, respectively. Compared to the period before the pandemic, excess deaths from all-causes, cardiovascular and respiratory diseases, pneumonia and influenza attributable to influenza A and B significantly decreased in all age groups. However, excess deaths associated with ozone increased in all age-disease categories, while the relative change of nitrogen dioxide (NO_2) and particular matters less than 10 μm (PM_{10}) associated burden showed a varied pattern.

Conclusions: A notable shift in disease burden attributable to influenza and air pollutants was observed in the pandemic period, suggesting that both direct and indirect impacts shall be considered when assessing the global and regional burden of the COVID-19 pandemic.

Keywords

machine learning, mortality, COVID-19, influenza, air pollution

Submission date: 28 October 2023; Acceptance date: 29 May 2024

Introduction

As of 10 March 2024, the World Health Organization (WHO) reported that the COVID-19 pandemic has caused 775 million infections and 7 million deaths worldwide.[1] Prior to the availability of vaccines in early 2021, global health authorities adopted various strategies to curb the spread of the virus. Mainland China and the Hong Kong Special Administrative Region, which implemented a Zero-COVID policy to eradicate community outbreaks, enforced stringent public health and social measures. These included strict border control, mandatory quarantine

[1]Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong SAR, China
[2]JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong SAR, China
[3]School of Nursing, Hong Kong Polytechnic University, Hong Kong SAR, China
[^]These authors contributed equally to this work.

Corresponding authors:
Lin Yang, GH515, School of Nursing, Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China.
Email: l.yang@polyu.edu.hk;

Daihai He, TU804, Department of Applied Mathematics, Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China.
Email: daihai.he@polyu.edu.hk.

for incoming travelers, and the promotion of social distancing through work-from-home policies and school closures.[2–4] Mask mandates were also enforced from July 2020.[5] As a result, Hong Kong recorded a relatively low number of COVID-19 cases and a low mortality rate in the first year of the pandemic, with only 148 deaths in 2020.[6]

The COVID-19 outbreak has significantly interfered the seasonal patterns of other respiratory viruses. Influenza activity, for instance, has remained low in many regions or countries since 2020, largely due to enhanced personal hygiene and social distancing measures.[7] The 2019–2020 winter influenza season in Hong Kong was remarkably shorter than previous seasons,[8] which likely resulted in a significant reduction in disease burden attributable to the influenza virus. Prior to the pandemic, the WHO estimated that seasonal influenza was associated with 3 to 5 million cases globally and 290,000 to 650,000 respiratory deaths annually.[9] In Hong Kong, our previous studies estimated that influenza resulted in approximately one thousand annual deaths and tens of thousands of hospitalizations.[10,11] Most influenza-associated deaths occurred in people with chronic diseases and the elderly, who were also at higher risks of severe infections and deaths from COVID-19. Hence, the displacement of mortality or morbidity of susceptible populations by COVID-19 might have offset the disease burden caused by influenza. Similarly, improved air quality has been reported during the pandemic, likely due to social distancing and reduced economic activities.[12] While the direct effects of these containment measures on COVID-19 have been intensively evaluated,[3,13] few studies have investigated their indirect effects on the disease burden of other respiratory viruses and air pollution. A few studies estimated excess all-cause mortality in the pandemic across countries.[14] However, most just calculated the difference between observed deaths and seasonal baseline based on simple time-series models, few have considered other factors such as air pollution, temperature, and influenza in their models.

Previous ecological studies on the disease burden associated with influenza have often used time-series modeling approaches such as quasi-Poisson regression or linear regression models.[15,16] These models share a common approach to calculating excess mortality. Initially, a model is fitted to the observed data and seasonal baselines are estimated from the model under the assumption that virus proxies are zero. The disease burden is then quantified by excess mortality, calculated as the difference between the observed data and the predicted baseline data. However, the selection of regression models and virus proxies remains a topic of debate. For example, the quasi-Poisson model with a log-link function has been used for weekly counts of all-cause and cause-specific mortality, but it has faced criticism for assuming a log-linear relationship between virus proxies and outcomes. The Gaussian linear model, which assumes a linear relationship between virus proxies and outcome variables, often yields negative effect estimates for virus proxies. These time-series models incorporate covariates such as seasonal trends, temperature, humidity, and proxies for influenza virus activities.[15,17]

In this study, we used a machine learning modeling strategy, eXtreme Gradient Boosting (XGBoost), to estimate the disease burden attributable to influenza and environmental factors. The XGBoost model offers the flexibility of assessing the marginal effect of each variable and minimizes overfitting through regularization penalties, bootstrapping of samples and cross-validation. This model has been applied to various research topics, including prediction models for disease incidence and prognosis, as well as the mortality burden of air pollution.[18–20] We also compared the XGBoost model estimates with those from the general additive model (GAM) with a Gaussian link function, which has been used in previous studies on the disease burden of influenza.[15,21]

Method

Data

We obtained individual death records from 2014 to 2021 in Hong Kong from the Census and Statistics Department of the Hong Kong Special Administrative Region, China. These records were then aggregated to weekly death counts based on the International Classification of Diseases, Tenth Revision (ICD-10). As in our previous studies,[10,11] we considered the following categories: all-cause (ICD-10 codes A00-Z99), cardiovascular and respiratory diseases (CRD, ICD-10 codes I00-I99&J00-J99), and pneumonia and influenza (P&I, ICD-10 codes J10-J18). We further divided weekly death counts into six age groups: 0–19, 20–39, 40–64, 65–84, and 85+ years. These age groups were selected to be consistent with our previous research on influenza disease burden.[10] Population data of these age groups were obtained from the Census and Statistics Department of Hong Kong. Weekly age-specific population size was calculated from annual mid-year age-specific population using a LOESS smooth function. Influenza surveillance data and weekly COVID-19 death counts were retrieved from the Centre for Health Protection. Meteorological data, including daily temperature and relative humidity, were downloaded from the Hong Kong Observatory website. Air pollution data, including hourly concentrations of particular matter less than 10 μm in diameter (PM_{10}), ozone (O_3) and nitrogen dioxide (NO_2), were retrieved from the 18 General and Roadside Air Quality Monitoring Stations of the Environmental Protection Department.

Statistical models

Model development. We first trained the XGBoost models on weekly death data by age groups. Weekly proportions

of specimens testing positive for influenza type A (subtype H1N1 and H3N2) or type B were added to the models as proxy variables for influenza. The models also included weekly averages of mean, maximum, and minimum temperature, weekly average relative humidity, and weekly average concentration of PM_{10}, O_3 and NO_2. Other covariates in the models were week and year dummies to adjust for seasonal and annual trends. Weekly age-specific population was used as an offset in the model to adjust for population size.

Before fitting the models, weekly COVID-19 deaths were subtracted from weekly mortality data. We first trained the XGBoost model by repeatedly fitting models to the pre-pandemic data (2014–2019). We assumed different learning rates and maximum tree depth during the validation process with a fivefold cross-validation to prevent overfitting. The training procedure of the XGBoost models was conducted separately on all-cause mortality and cause-specific mortality by different age groups. The model parameters and number of boosted trees were decided based on the log-likelihood.

We also adopted a classical time-series modeling approach, the GAM model with a Gaussian link function, which has been used in previous studies on the disease burden of influenza and air pollution. We fit the GAM models to all-cause mortality and cause-specific mortality, with adjustment for overfitting.[22] The best-fit GAM model was selected after fivefold cross-validation. Cubic splines were applied to covariates like week number, average temperature, and relative humidity with 10 knots each. Similar to the XGBoost model, the GAM models also added weekly age-specific population size as an offset to adjust for age structure change over time.

Estimation of mortality burden. The mortality burden during the pre-pandemic and pandemic period was assessed using the following measurements: (1) *Overall excess mortality in the pandemic:* this was calculated as the difference between the observed death counts in 2020–2021 and the predicted counts from the XGBoost and GAM models using the observed data of covariates (meteorological data, air pollutants, and influenza). (2) *Excess mortality associated with influenza:* We first estimated the *influenza baseline mortality* by setting the influenza proxy variable in 2020–2021 to zero (i.e. assuming no influenza virus activity) in the models. The difference between the observed death counts in 2020–2021 and the influenza baseline mortality was the excess mortality specifically attributable to influenza. (3) *Excess mortality associated with air pollution:* Using PM_{10} as an example, we first estimated the *PM_{10} baseline mortality* by setting the PM_{10} variable in 2020–2021 to its minimum value during the pre-pandemic period (9.7 μg/m^3) and subtracted this baseline data from the observed data to calculate excess mortality specifically attributable to PM_{10}. A similar calculation was repeated for NO_2 and O_3 (minimum values 23.8 and 13.8 μg/m^3, respectively). Considering the potential delay in mortality effects, we also calculated the lag effects up to 14 days prior for each influenza proxy and air pollutant. The 95% confidence interval (CI) for each estimate was calculated using bootstrapping for 10,000 times.

All data analysis was conducted using the "H2O" and "xgboost" packages in R software (version 4.1.0) (R Foundation for Statistical Computing, Vienna, Austria). The R codes are publicly accessible at https://github.com/yanglinpolyu/covid-excess-mortality.

Results

Mortality burden during the pre-pandemic and pandemic period

Figure 1 shows a comparison of weekly mortality rates specific to different age groups during the pre-pandemic (2014–2019) and pandemic period (2020–2021). Mortality rates for all categories increased in the age groups of 20–39, 40–64, and 85+ during the pandemic compared to the pre-pandemic period. The numbers of all-cause deaths showed a sudden drop from February to April 2020, but gradually increased thereafter, peaking in early 2021 (Figure 2). Influenza nearly disappeared during these two years, with a small peak of influenza type B in winter 2020 only (Figure S1). Air pollutants generally remained at a relatively low level during the pandemic, with a sudden drop in early 2020 (except O_3), and gradually returned to pre-pandemic levels thereafter.

Overall excess mortality during the pandemic

Both the XGBoost and GAM models fitted well to the weekly data of all-cause mortality counts from 2014 to 2019 (Figure 2). The XGBoost models outperformed the GAM models in terms of Model goodness-of-fit and prediction accuracy (Table S1). Given its capability of handling non-linear relationships and multicollinearity of predictors (Figure S2), we adopted the XGBoost models for main analysis and reported the estimates from these models hereafter. The greatest increase in excess mortality was observed in the 65–84 and 85 + age groups, with little to no significant increase in children and adolescents (Figure 2). A similar change in excess mortality was found in CRD, but not in P&I (Figures S3 and S4). There was a significant increase in all-cause excess mortality in Hong Kong during the first 2 years of the COVID-19 pandemic, with an estimate of 8762 (95% CI, 7503–9993), and 12,496 (95% CI, 11,718–13,332) in 2020 and 2021, respectively (Table 1). These correspond to 117.4 and 167.9 per 100,000 population, 12.6% and 8.5% of total deaths in these years. The highest rate of excess mortality

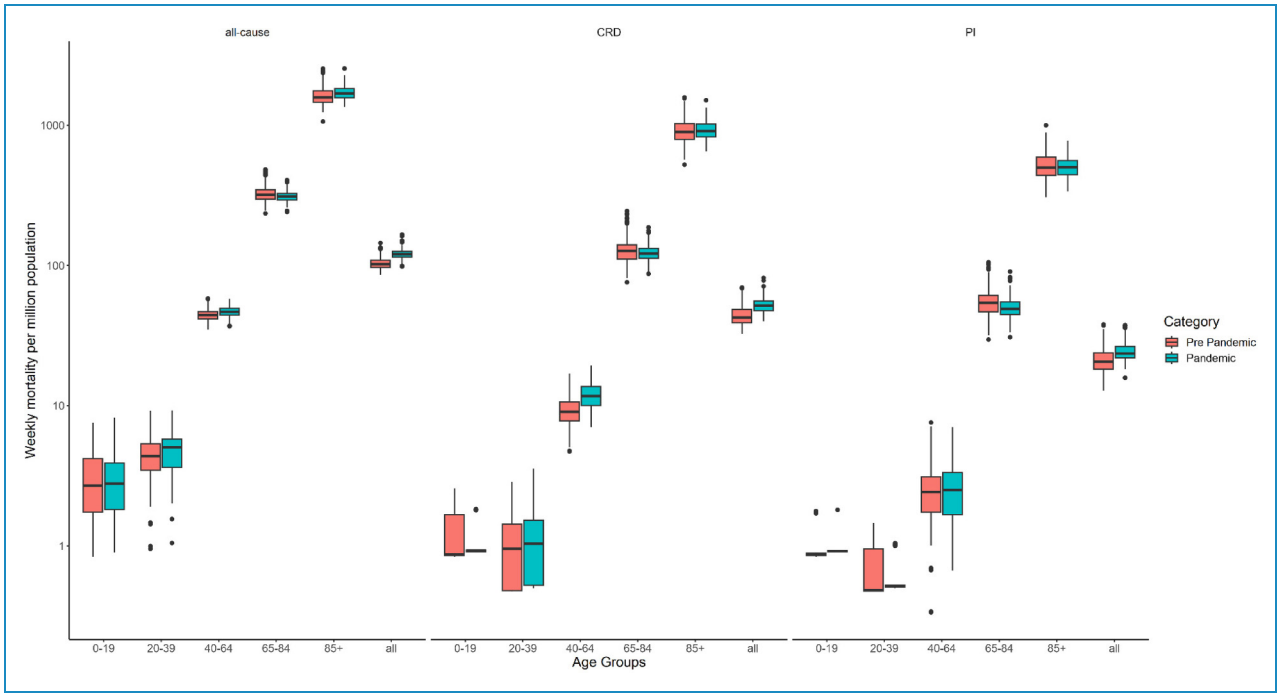

Figure 1. Weekly all-cause and cause-specific mortality by age group, during the pre- and pandemic period. CRD, cardiovascular and respiratory diseases; P&I, pneumonia and influenza.

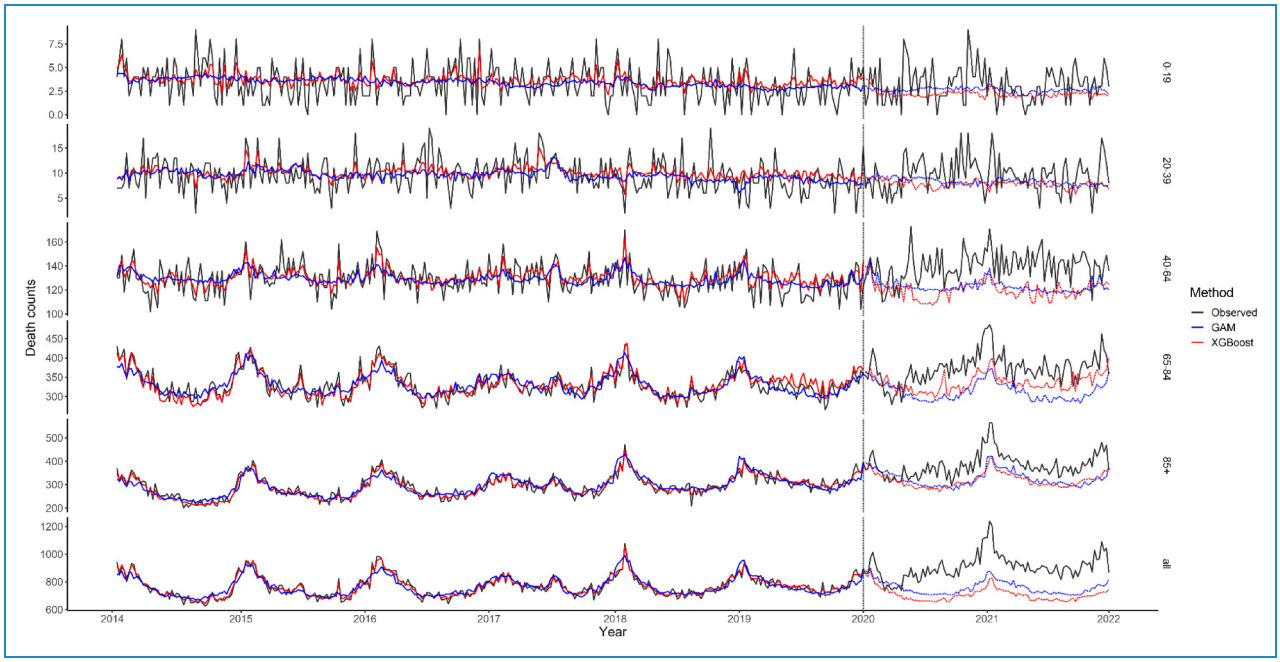

Figure 2. Observed and fitted weekly all-cause mortality data by age groups. The XGBoost (blue line) and GAM models (red line) were developed from the training data from 2014 to 2019 and used to predict the data in 2020 and 2021. XGBoost: eXtreme Gradient Boosting; GAM: general additive model.

was found in those aged 65–84 years, with the excess rate estimates of 1263.4 and 1797.1 per 100,000 population, respectively (Table 1).

Excess mortality associated with influenza

During 2014–2019, influenza A (H1N1 and H3N2) resulted in more deaths than influenza B in children and adolescents,

Table 1. Overall excess mortality numbers and rates (per 100,000 population) of all-causes, cardiovascular diseases (CRD), pneumonia and influenza (P&I) by age group in 2020 and 2021, estimated from the XGBoost models, respectively.

	All-cause				CRD				P&I			
	Excess number	(95% CI)	Excess rate	(95% CI)	Excess number	(95% CI)	Excess rate	(95% CI)	Excess number	(95% CI)	Excess rate	(95% CI)
2020												
All age	8762	(7503, 9993)	117.36	(100.17,134.71)	4221	(3571, 4874)	56.59	(47.82,65.36)	2199	(1874, 2530)	29.52	(25.22,34.08)
0–19	49	(19, 80)	4.48	(1.76,7.27)	−1	(-9, 9)	−0.07	(−0.82,0.79)	−1	(−6, 6)	−0.08	(−0.55,0.5)
20–39	108	(57, 159)	5.49	(2.9,8.07)	42	(19, 66)	2.12	(0.96,3.32)	1	(−7, 10)	0.06	(−0.36,0.52)
40–64	805	(590, 1018)	27.03	(19.63,34.14)	677	(557, 800)	22.63	(18.6,26.71)	123	(79, 169)	4.13	(2.64,5.67)
65–84	1399	(888, 1889)	118.65	(75.26,160.07)	989	(668, 1317)	84.31	(55.62,112.52)	431	(279, 587)	36.80	(23.45,50.62)
85+	2750	(2123, 3413)	1263.36	(978.04,1562.29)	960	(617, 1300)	440.54	(279.09,598.14)	672	(465, 887)	311.33	(214.32,405.93)
2021												
All age	12496	(11718, 13332)	167.86	(157.43,179.48)	6008	(5539, 6534)	80.60	(74.29,87.56)	2911	(2675, 3150)	39.07	(35.93,42.27)
0–19	29	(8, 50)	2.66	(0.74,4.63)	−5	(−11, 1)	−0.47	(−1.04,0.12)	0	(−4, 5)	0.04	(−0.33,0.49)
20–39	111	(66, 156)	5.75	(3.5,8.01)	37	(21, 54)	1.93	(1.07,2.84)	2	(−6, 11)	0.12	(−0.32,0.59)
40–64	1150	(976, 1325)	38.52	(32.7,44.17)	811	(699, 919)	27.02	(23.38,30.65)	103	(59, 147)	3.43	(1.96,4.91)
65–84	1770	(1325, 2221)	146.33	(109.4,182.39)	1312	(1086, 1552)	107.96	(88.82,127)	412	(271, 546)	33.87	(22.47,45.15)
85+	4062	(3650, 4480)	1797.08	(1618.03,1982.17)	1518	(1226, 1815)	672.23	(544.9,803.74)	973	(792, 1165)	431.59	(346.89,512.62)

XGBoost: eXtreme Gradient Boosting; CI: confidence interval.

but fewer in older adults (Figure 3). A similar pattern was observed in CRD and P&I, except for more P&I deaths associated with influenza A in the 85 + group (Figures S5 and S6). Annual excess mortality rates attributable to influenza A and B dramatically decreased in all age groups during the pandemic period, compared to the pre-pandemic period (Figure 4). The largest decrease (87% reduction from pre-pandemic estimates) was observed in excess mortality associated with PM_{10} in the 0–19 group, and the highest increase (273.47 times higher) was observed in influenza B-associated deaths in the 85+ group. Similarly, annual rates of CRD and P&I excess mortality attributable to influenza A and B dramatically decreased in all age groups during the pandemic (Figures S7 and S8).

Excess mortality associated with air pollution

During the pre-pandemic period, annual all-cause excess mortality associated with NO_2 and PM_{10} showed a decreasing trend across years in all age groups, while O_3 estimates

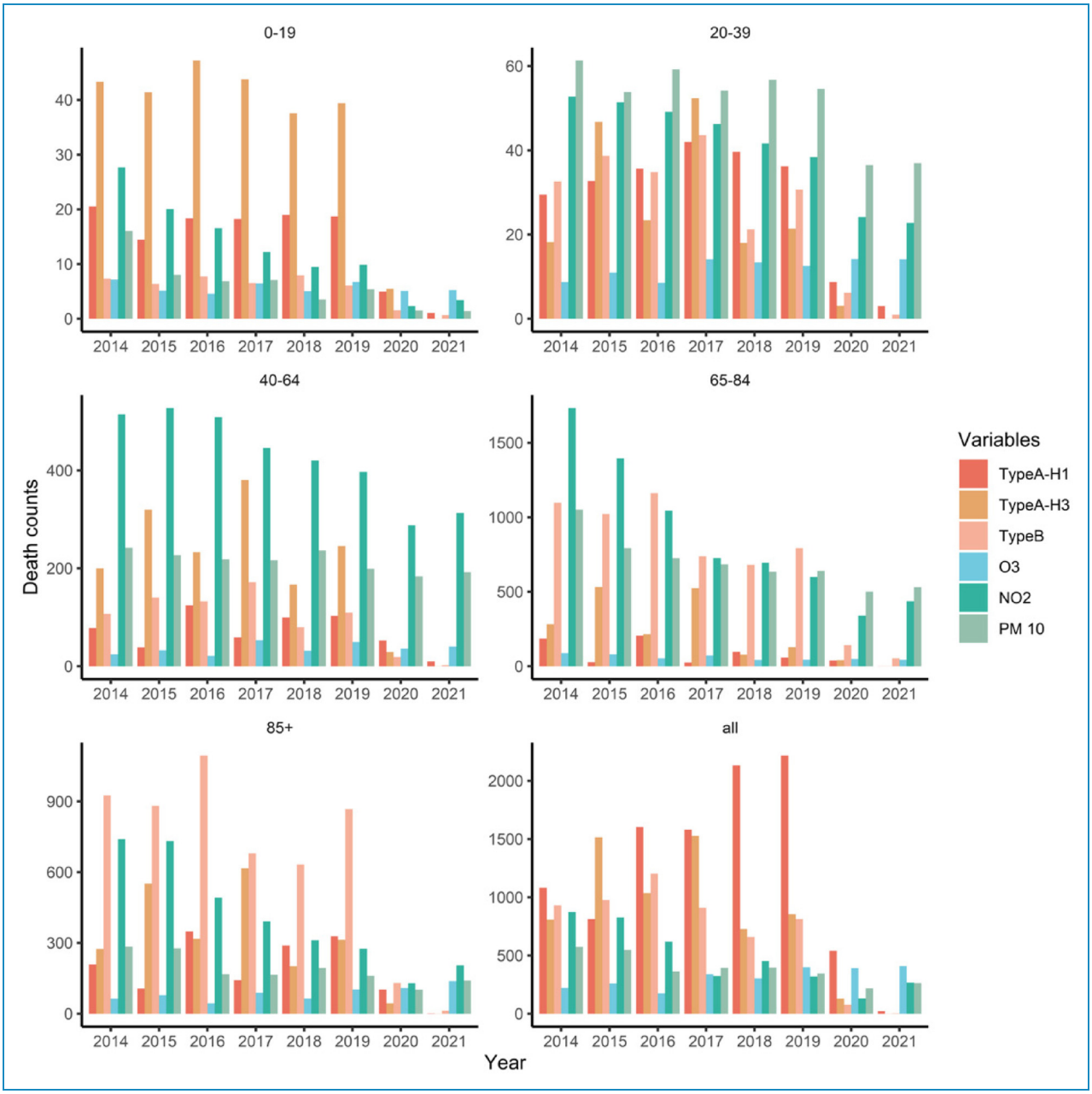

Figure 3. Annual all-cause mortality associated with influenza A subtype H1N1 (red bar), H3N2 (orange bar), influenza B (pink bar), O_3 (blue bar), NO_2 (green bar), and PM_{10} (light green bar). The estimates were derived from the XGBoost model. XGBoost: eXtreme Gradient Boosting.

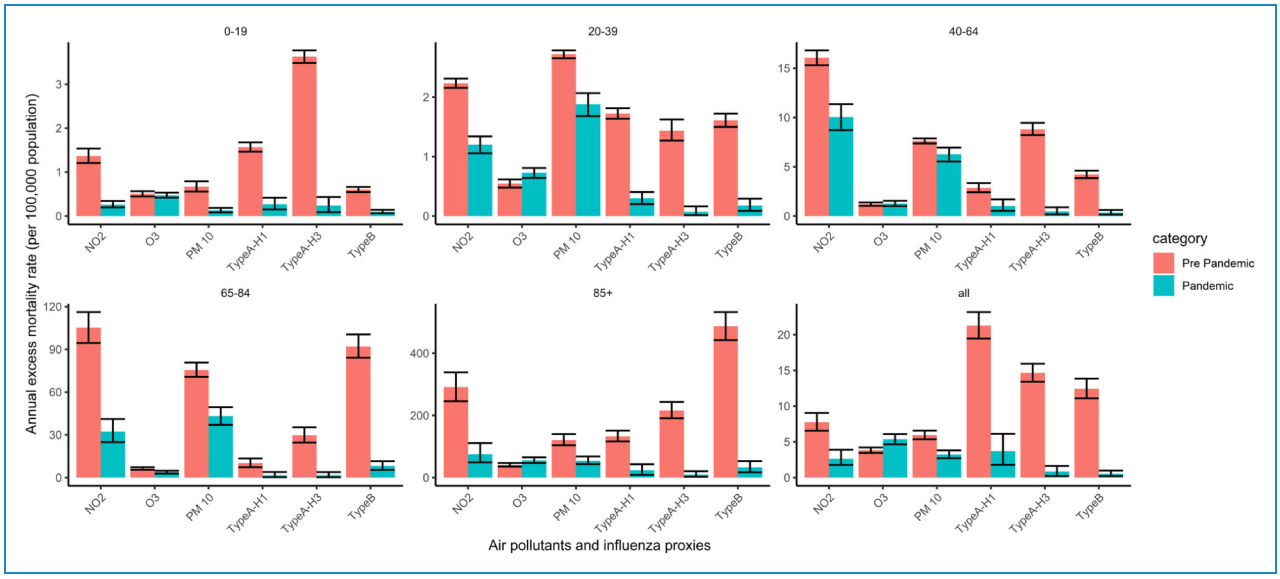

Figure 4. Comparison of annual all-cause excess mortality rate (per 100,000 population) associated with air pollutants, influenza A and B between the pre- and pandemic period, for different age groups, with 95% confidence interval (error bar) obtained from bootstrapping for 10,000 times.

slightly increased (Figure 3). Annual all-cause excess mortality rates attributable to air pollutants decreased in all age groups during the pandemic period, with a few exceptions including O_3 estimates in the all age group (Figure 4). The estimates of CRD and P&I excess mortality rates attributable to all air pollutants had a similar pattern (Figures S7 and S8).

We also conducted a sensitivity analysis by incorporating data on air pollutants and influenza proxies from up to 14 days before (Figure S9). The model estimates are generally similar to the main results. Compared to the XGBoost models, the estimates from the GAM models were more conservative than those from showing negative values in several age-disease categories. However, the overall trends of relative change remained consistent between the two models (Table S2).

Discussion

In this study, we observed a significant rise in all-cause, CRD and P&I mortality during the pandemic in Hong Kong. The XGBoost model estimated excess mortality at 21,258, 10,229, and 5110, while the GAM model estimated it at 15,431, 8426, and 2666, respectively. These numbers are significantly higher than the 13 COVID-19 deaths reported during the same period. As expected, older age groups (65–84 and 85+ years) had a greater increase than younger age groups. Interestingly, the annual all-cause excess mortality rates attributable to both influenza and three common air pollutants decreased in all age groups during the pandemic compared to the pre-pandemic period, with the exception of a slight increase in all-cause excess mortality rates associated with O_3. Although we included time, temperature, relatively humidity, respiratory viruses, and air pollution as covariates in the models, there were likely other unadjusted factors contributing to the increased mortality burden. We hypothesize that changes in health-seeking behavior could be one potential reason, with people possibly delaying diagnosis and treatment due to fears of contracting COVID-19 in healthcare facilities. This is supported by reports of a significant decrease in overall hospitalizations in Hong Kong in 2020, with a corresponding decrease in in-hospital mortality and an increase in out-of-hospital mortality.[23]

Most previous studies on excess mortality only considered time and seasonal trends in predicting baseline deaths, and some even produced negative estimates.[24] This highlights the difficulty in assessing the true impact of the pandemic. Our study is one of the first to conduct a comprehensive investigation into the relative change in disease burden associated with different factors, including influenza and air pollutants. It is not surprising to observe a dramatic drop in all-cause excess mortality rates attributable to both influenza A and influenza B during the pandemic. This is likely due to enhanced personal hygiene and social distancing measures, which have been reported to have nearly eliminated other respiratory pathogens during the COVID-19 pandemic in many countries and regions.[25,26] Additionally, the uptake rates of seasonal influenza vaccination increased in the Hong Kong population, particularly among children aged 6 months to 6 years, with rates increasing from 19.2% in 2016/17 season to 47.4% in 2019/20 season.[27]

Previous studies have investigated changes in air pollutants during lockdowns. In China, it was found that NO_2 and CO concentrations decreased during the lockdown period, while O_3 increased. Based on these three air pollutants, the lockdown policy prevented certain all-cause deaths.[28] A global study involving air quality stations from 34 countries estimated that a net total of 49,900 excess deaths were avoided during lockdowns due to reduced emission of NO_2, O_3 and $PM_{2.5}$. In China, the $PM_{2.5}$-related avoided excess mortality was 19,600.[29] However, studies in England and Australia found no association between air pollution and excess mortality when comparing the pre-pandemic and pandemic periods,[30] and the change in air quality during the COVID-19 lockdown had a negligible impact on calculated health outcomes.[31] Whether the reduction in air pollution during the lockdown has an effect on total mortality remains a question, but studies have found that air quality is the most important factor in the context of enabling an increase in COVID-19 case load.[32] A review by Becchetti et al. (2021) found a strong association between long-term air pollution exposure and COVID-19 deaths.[33] In our study, we found that the annual all-cause, CRD, and P&I excess deaths attributed to NO_2 and PM_{10} decreased during the pandemic in Hong Kong based on our XGBoost model, while the excess deaths attributed to O_3 increased slightly, similar to the study in China.[28] We also observed an increase in the excess deaths attributed to O_3 in the pre-pandemic period. In addition to the excess deaths attributed to air pollutants, we found a significant decrease in the annual all-cause, CRD, and P&I excess deaths attributed to both influenza type A and type B. Although the predicted all-cause weekly mortality counts for each age group during the pandemic period were lower than those in the pre-pandemic period due to the decrease in excess deaths attributed to major air pollutants and influenza proxies (with the exception of excess deaths attributed to O_3), there was a significant amount of all-cause and CRD excess mortality counts for most age groups based on both prediction methods after deducting the weekly COVID-19 deaths.

Few studies have simultaneously assessed the disease burden of influenza and air pollutants, primarily due to the multicollinearity between these variables (Figure S2). This challenge is effectively managed by our XGBoost models, which utilize decision trees and regularization techniques to deal with non-linear relationships. While accurately determining the disease burden from influenza and air pollution is challenging, our calculated excess mortality for the period before the pandemic aligns well with findings from earlier research. For instance, a study conducted in China in 2017 estimated around 4200 deaths (with a range of 3300–5100) in Hong Kong was due to air pollution.[34] Another study estimated around 3000 deaths in Hong Kong in 2013 were attributable to air pollutants.[35] Regarding the burden from influenza, estimates from previous studies using GAM models suggested between 500 and 1000 excess deaths from influenza A and B in Hong Kong between 1998 and 2009.[10,15] These numbers are within the range of our XGBoost model estimates, which suggest between 4500 and 5000 excess deaths from 2014 to 2019 due to these factors (Figure 3). Taken together, these findings support the reliability of the XGBoost models in providing estimates for the disease burden from air pollutants and influenza.

There are some limitations in this study. First, we only retrieved meteorological data, air pollution data, and influenza data. We did not include any demographic information. Second, we chose the GAM model with a Gaussian link function, which may have negative effect estimates of influenza, making it difficult to compare the influenza and air pollution-associated mortality burden calculated with different models. Finally, further research should be conducted to explore the possible reasons for excess mortality in Hong Kong during the pandemic period.

Conclusion

Using advanced machine learning approaches, we estimated a significant decrease in disease burden associated with influenza and air pollutants in a region with minimal COVID-19 cases in the first 2 years of the pandemic, while the overall mortality burden during the pandemic period increased compared to the pre-pandemic period. Our findings suggest that comprehensive assessments of the global and regional burden associated with the COVID-19 pandemic should consider its direct and indirect impacts.

Acknowledgements: The authors sincerely thank the Dr Peihua Cao for his interest and support of this project.

Contributorship: YL, JX, and CC wrote the first draft of the manuscript. All other authors contributed to revising the manuscript.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by Health and Medical Research Fund (HMRF) commissioned studies on COVID-19 research (reference number COVID1903007).

Ethical approval: Ethical approval was not required since this study did not involve human participants.

Informed consent: No patient consent was required for this study. All data used in this study were obtained from a database that guaranteed complete anonymity.

Guarantor: LY is the guarantor.

ORCID iD: Lin Yang https://orcid.org/0000-0002-5964-3233

Supplemental material: Supplemental material for this article is available online.

References

1. WHO. Coronavirus disease (COVID-19) pandemic. 2020.
2. Lai CK, Ng RW, Wong MC, et al. Epidemiological characteristics of the first 100 cases of coronavirus disease 2019 (COVID-19) in Hong Kong special administrative region, China, a city with a stringent containment policy. *Int J Epidemiol* 2020; 49: 1096–1105.
3. Wong MC, Ng RW, Chong KC, et al. Stringent containment measures without complete city lockdown to achieve low incidence and mortality across two waves of COVID-19 in Hong Kong. *BMJ global Health* 2020; 5: e003573.
4. Koh WC, Alikhan MF, Koh D, et al. Containing COVID-19: implementation of early and moderately stringent social distancing measures can prevent the need for large-scale lockdowns. *Ann Glob Health* 2020; 86: 88.
5. SAR HK. Government further tightens social distancing measures, 2020. https://www.info.gov.hk/gia/general/202007/27/P2020072700650.htm?fontSize=1.
6. Health CfHPotDo. Archive of Statistics on Provisional Analysis on Reported Death Cases, https://www.coronavirus.gov.hk/eng/death_analysis.html (2022, 2023).
7. Jones N. How COVID-19 is changing the cold and flu season. *Nature* 2020; 588: 388–390.
8. Chan KH, Lee PW, Chan CY, et al. Monitoring respiratory infections in COVID-19 epidemics. *Br Med J* 2020; 369: m1628.
9. Organization WH. Influenza (Seasonal), 2018, https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal).
10. Yang L, Chan K, Cowling B, et al. Excess mortality associated with the 2009 pandemic of influenza A (H1N1) in Hong Kong. *Epidemiology & Infection* 2012; 140: 1542–1550.
11. Yang L, Wang X, Chan K, et al. Hospitalisation associated with the 2009 H1N1 pandemic and seasonal influenza in Hong Kong, 2005 to 2010. *Eurosurveillance* 2012; 17: 20121108.
12. Venter ZS, Aunan K, Chowdhury S, et al. COVID-19 lockdowns cause global air pollution declines. *Proc Natl Acad Sci USA* 2020; 117: 18984–18990.
13. Signorelli C, Scognamiglio T and Odone A. COVID-19 in Italy: impact of containment measures and prevalence estimates of infection in the general population. *Acta Bio Medica: Atenei Parmensis* 2020; 91: 175.
14. Mathieu E, Ritchie H, Rodés-Guirao L, et al. Coronavirus pandemic (COVID-19). Our world in data 2020. https://ourworldindata.org/coronavirus (2023, accessed 1 May 2023).
15. Wu P, Goldstein E, Ho LM, et al. Excess mortality associated with influenza A and B virus in Hong Kong, 1998–2009. *J Infect Dis* 2012; 206: 1862–1871.
16. Wang XL, Yang L, Chan KP, et al. Model selection in time series studies of influenza-associated mortality. *PLoS One* 2012; 7: e39423.
17. Yu X, Wang C, Chen T, et al. Excess pneumonia and influenza mortality attributable to seasonal influenza in subtropical Shanghai, China. *BMC Infect Dis* 2017; 17: 1–9.
18. Desdhanty VS and Rustam Z. Liver cancer classification using random forest and extreme gradient boosting (xgboost) with genetic algorithm as feature selection. In: *2021 International conference on decision aid sciences and application (DASA)*. IEEE, New York City, United States; 2021, pp.716–719.
19. Dhillon S, Bansal C and Sidhu B. Machine learning based approach using XGboost for heart stroke prediction. In: *International conference on emerging technologies: AI, IoT, and CPS for science & technology applications.* NITTTR Chandigarh, India; 2021.
20. Ding J, Liu M, Ma Z, et al. Spatial and temporal trends in the mortality burden of ozone pollution in China: 2005-2017. In: *ISEE conference abstracts*. Volume 2020, Issue 1. EHP Publishing, Durham, North Carolina, United States.
21. Wong JY, Wu P, Nishiura H, et al. Infection fatality risk of the pandemic A (H1N1) 2009 virus in Hong Kong. *Am J Epidemiol* 2013; 177: 834–840.
22. Hilt DE and Seegrist DW. *Ridge, a computer program for calculating ridge regression estimates*. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, 1977. Upper Darby, PA, United States.
23. Xin H, Wu P, Wong JY, et al. Hospitalizations and mortality during the first year of the COVID-19 pandemic in Hong Kong, China: an observational study. *The Lancet Regional Health–Western Pacific* 2023; 30: 100645.
24. Beaney T, Clarke JM, Jain V, et al. Excess mortality: the gold standard in measuring the impact of COVID-19 worldwide? *J R Soc Med* 2020; 113: 329–334.
25. Chiu NC, Chi H, Tai YL, et al. Impact of wearing masks, hand hygiene, and social distancing on influenza, enterovirus, and all-cause pneumonia during the coronavirus pandemic: retrospective national epidemiological surveillance study. *J Med Internet Res* 2020; 22: e21257.
26. Chow EJ, Uyeki TM and Chu HY. The effects of the COVID-19 pandemic on community respiratory virus activity. *Nat Rev Microbiol* 2023; 21: 195–210.
27. SAR HK. Statistics on seasonal influenza vaccination, 2020. https://www.info.gov.hk/gia/general/202011/18/P2020111800593.htm?fontSize=1.
28. Xu Z, Cao R, Hu X, et al. The improvement of air quality and associated mortality during the COVID-19 lockdown in one megacity of China: an empirical strategy. *Int J Environ Res Public Health* 2021; 18: 8702.
29. Venter ZS, Aunan K, Chowdhury S, et al. Air pollution declines during COVID-19 lockdowns mitigate the global health burden. *Environ Res* 2021; 192: 110403.

30. Davies B, Parkes BL, Bennett J, et al. Community factors and excess mortality in first wave of the COVID-19 pandemic in England. *Nat Commun* 2021; 12: 3755.
31. Ryan RG, Silver JD and Schofield R. Air quality and health impact of 2019–20 black summer megafires and COVID-19 lockdown in Melbourne and Sydney, Australia. *Environ Pollut* 2021; 274: 116498.
32. Kowalski PA, Szwagrzyk M, Kielpinska J, et al. Numerical analysis of factors, pace and intensity of the corona virus (COVID-19) epidemic in Poland. *Ecol Inform* 2021; 63: 101284.
33. Becchetti L, Beccari G, Conzo G, et al. Air quality and COVID-19 adverse outcomes: divergent views and experimental findings. *Environ Res* 2021; 193: 110556.
34. Yin P, Brauer M, Cohen AJ, et al. The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990-2017: an analysis for the global burden of disease study 2017. *Lancet Planet Health* 2020; 4: E386–E398.
35. Lu XC, Yao T, Fung JCH, et al. Estimation of health and economic costs of air pollution over the Pearl River Delta region in China. *Sci Total Environ* 2016; 566: 134–143.