

Socioeconomic Influences on COVID-19 Incidence in Hong Kong Public Housing: Insights for Future Public Health Strategies

John David Pilapil

jdpilapil@nus.edu.sg

National University Singapore Saw Swee Hock School of Public Health <https://orcid.org/0000-0002-0767-2454>

Jean WOO

The Chinese University of Hong Kong

Mei-Po KWAN

The Chinese University of Hong Kong

Angkana HUANG

National University of Singapore

King Lun YEUNG

The Hong Kong University of Science and Technology

Research Article

Keywords: Health inequity, Public housing, Social determinants of health, Geospatial epidemiology, Hong Kong, COVID-19

Posted Date: October 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11112167/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Socioeconomic Influences on COVID-19 Incidence in Hong Kong Public Housing: Insights for Future Public Health Strategies

John David PILAPIL^{1,2}, Jean WOO³, Mei Po KWAN⁴, Angkana HUANG², King Lun

YEUNG^{2,5,6,*}

¹Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China

²Saw Swee Hock School of Public Health, National University of Singapore, Science Drive, Singapore

³Department of Medicine & Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

⁴Department of Geography and Resource Management, and Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

⁵Division of Environment and Sustainability, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China

⁶Thrust of Earth, Ocean, and Atmospheric Sciences of Function Hub, The Hong Kong University of Science and Technology (Guangzhou), 1 Du Xue Rd. Nansha District, Guangzhou, China

**Corresponding Author:*

King Lun Yeung

Chair Professor, The Hong Kong University of Science and Technology

Clear Water Bay, Kowloon, Hong Kong SAR

kekyeung@ust.hk

Abstract

Socioeconomic vulnerabilities inherent in public housing can disproportionately amplify risks from infectious diseases. This work presents a cross-sectional study of cumulative COVID-19 incidence (2020-2022) across 197 Hong Kong public housing estates, integrating surveillance data with demographic and socioeconomic profiles. Geographic disparities assessed by Kruskal-Wallis tests (η^2 effect sizes) were significant at regional ($H = 20.92$, $p < 0.001$, $\eta^2 = 0.107$) and district levels ($H = 45.67$, $p < 0.001$, $\eta^2 = 0.233$), with non-random spatial clustering confirmed by Global Moran's I ($I = 0.194$, $p = 0.002$). Incidence varied 340-fold across estates (0.04-13.64 per 1,000). After Spearman's rank correlation with Benjamini-Hochberg FDR correction of socioeconomic variables, no predictor achieved $q < 0.05$; the strongest nominally significant correlates were average household size ($\rho = -0.192$), rent-to-income ratio ($\rho = 0.182$), and single-person household proportion ($\rho = 0.173$). A hyperparameter-tuned Random Forest regression model explained 22.5% of the variance (test $R^2 = 0.225$; $RMSE = 0.364$), with SHAP analysis identifying income poverty, the elderly isolation rate, and the proportion of craft workers as the principal structural drivers. Therefore, COVID-19 burden in Hong Kong's public housing reflects spatially clustered, quantifiable socioeconomic disparities, in which income poverty, elderly isolation, and occupational exposure represent modifiable preparedness priorities.

Keywords: Health inequity, Public housing, Social determinants of health, Geospatial epidemiology, Hong Kong, COVID-19

1. INTRODUCTION

Urban public housing estates are most vulnerable to infectious disease as they concentrate fundamental drivers of health inequity: poverty, precarious employment, substandard housing, and restricted access to healthcare [1–4]. The COVID-19 pandemic made these inequities visible at scale. Across high-income countries, low-income urban communities experienced disproportionately high burdens of infection, hospitalization, and mortality, driven by overcrowded housing, a high proportion of essential workers, and limited capacity to adopt protective behaviors [5–7]. Hong Kong exemplifies this with a population density exceeding 6,300 persons/km² and approximately 45% of its 7.5 million residents housed in estates managed by the Hong Kong Housing Authority [8]. During the COVID-19 pandemic, the city experienced a surge in infections (30-40% of the population) with the introduction of the more virulent Omicron BA.2 sub-variant despite initial success [9,10].

Existing studies on COVID-19 in Hong Kong's public housing have focused on architectural and environmental determinants, with building morphology, ventilation systems, and floor-level design associated with infection risk [11–14]. Citywide and regional modeling have investigated epidemic-wave dynamics and the effects of non-pharmaceutical interventions [15,16]. District-level analyses examined inequities in healthcare utilization and social monitoring frameworks [17,18]. However, there is a lack of systematic examination of the extent to which household-level socioeconomic characteristics, including income, housing cost burden, occupational class, and household composition, drive differential COVID-19 incidence across individual public housing estates, which are critical for developing a targeted, equity-oriented allocation of resources.

This study employs a multi-method estate-level analysis of cumulative COVID-19 incidence across all 197 major public housing estates in Hong Kong (2020-2022), grounded in

the Dahlgren-Whitehead social determinants of health model. We pursue three objectives: (1) characterize the geographic distribution of estate-level incidence using geospatial mapping and spatial autocorrelation; (2) identify socioeconomic correlates of estate-level risk using Spearman's rank correlation with Benjamini-Hochberg FDR correction; and (3) develop and interpret a Random Forest regression model with SHAP global explainability and LIME estate-level transparency to quantify the structural drivers of COVID-19 risk.

2. METHODS

2.1. Study design and setting

A cross-sectional study of cumulative 2020-2022 COVID-19 incidence on 197 major public housing estates operated by the Hong Kong Housing Authority (HKHA). The housing estate was considered as the unit of analysis, representing the smallest geographic unit for which both epidemiological and sociodemographic data are available. It is also the lowest administratively actionable level for HKHA and the Centre for Health Protection (CHP). As the data were aggregated across epidemic waves, temporal variation in circulating variants, non-pharmaceutical interventions, and vaccination coverage is an acknowledged limitation.

2.2. Data sources and variables

CHP COVID-19 case data provided confirmed cases for each public housing estate, and the cumulative incidence rates, expressed as cases per 1,000 residents, were calculated using estate-level population denominators [19]. Public housing characteristics for each estate were obtained from the HKHA, and the geographic coordinates provided distances between estates [20]. The full set of variables initially comprised approximately 38 estate-level predictors, spanning demographic composition, socioeconomic status and education, employment and occupational structure, living environment, and housing affordability (**Supplementary Figure**

1). It was narrowed to nine following multicollinearity audits, Spearman correlation analyses, and theoretical grounding.

2.3. Geospatial analysis

Estate-level COVID-19 incidence was visualized using choropleth and proportional-symbol maps. Regional and district-level disparities were assessed using the Kruskal-Wallis H test after non-normality had been confirmed by the Shapiro-Wilk test ($p < 0.05$). Effect size was quantified using eta-squared (η^2), and pairwise comparisons were performed using Dunn's test with Bonferroni correction. Global Moran's I was used to assess spatial autocorrelation in estate-level incidence, while local Moran's I (LISA) was used to identify significant high-high and low-low clusters. Analyses were conducted using GeoPandas (version 0.14.3), PySAL, and Matplotlib (version 3.8.4).

2.4. Socioeconomic correlation analysis

Spearman's rank correlation coefficient (ρ) was used to assess associations between each estate-level predictor and COVID-19 incidence. To account for multiple testing, p-values were adjusted using the Benjamini-Hochberg procedure to control the false discovery rate at 5% [21]. Both unadjusted and adjusted p-values are reported. Correlation strengths were interpreted as negligible ($|\rho| < 0.10$), weak (0.10-0.19), moderate (0.20-0.29), or strong (≥ 0.30). Variables were grouped into five thematic domains: demographic composition; socioeconomic status and education; employment and occupational structure; living environment; and housing affordability (**Figure S1**).

2.5. Random Forest, SHAP, and LIME

A Random Forest (RF) regression model was developed to predict log-transformed estate-level COVID-19 incidence rate, capturing non-linear and interactive effects among predictors. Multicollinearity was assessed using pairwise Spearman correlations and variance inflation

factors (VIFs); the predictor with $VIF = \infty$ and $|\rho| = 1.000$ relative to its composite index counterpart was excluded. The dataset was split into training and test sets in an 80:20 ratio ($n = 157$ and $n = 40$, respectively) using stratified random sampling. Hyperparameters were tuned using a five-fold cross-validated randomized search ($n_iter = 50$). Model performance was evaluated using root mean squared error (RMSE), mean absolute error (MAE), and R^2 on both cross-validation and held-out test sets.

Feature importance was assessed using mean decrease in impurity (MDI) and permutation importance. Global explainability was quantified using SHAP TreeExplainer [22], expressed as mean $|\phi|$ with 95% bootstrap confidence intervals ($n_boot = 1,000$), and visualized using beeswarm plots. Local explainability for purposively selected high-risk, intermediate-risk, and low-risk estates was examined using LIME (4,000 perturbations per estate) [23]. All analyses were conducted using Python 3.11, with Scikit-learn 1.4.2, SHAP 0.44.1, and LIME 0.2.0.1.

2.6. Ethical considerations

This study used exclusively anonymized, publicly available secondary data, including CHP COVID-19 surveillance records and HKHA estate-level profiles [19,20]. No individual patient data were accessed. In accordance with host institutional guidelines, formal ethical review was not required. All data handling complied with Hong Kong's Personal Data (Privacy) Ordinance (Cap. 486).

3. RESULTS

3.1. Geographic distribution and spatial patterning of COVID-19

Cumulative COVID-19 incidence across 197 public housing estates ranged from 0.04 to 13.64 cases per 1,000 residents ($mean = 1.73 \pm 1.53$), a more than 340-fold variation, indicating a highly heterogeneous distribution of disease burden (**Figures 1 and 2**). The ten estates with the highest burden include Kwai Chung Estate (13.64 per 1,000), Choi Fai Estate (8.92 per

1,000), Tsz Hong Estate (6.78 per 1,000), Fung Wah Estate (6.06 per 1,000), Ping Shek Estate (6.01 per 1,000), and Tsz Man Estate (5.90 per 1,000), while the lowest burden was recorded in Shek Lei (2) Estate (0.04 per 1,000) and Choi Wan (1) Estate (0.06 per 1,000).

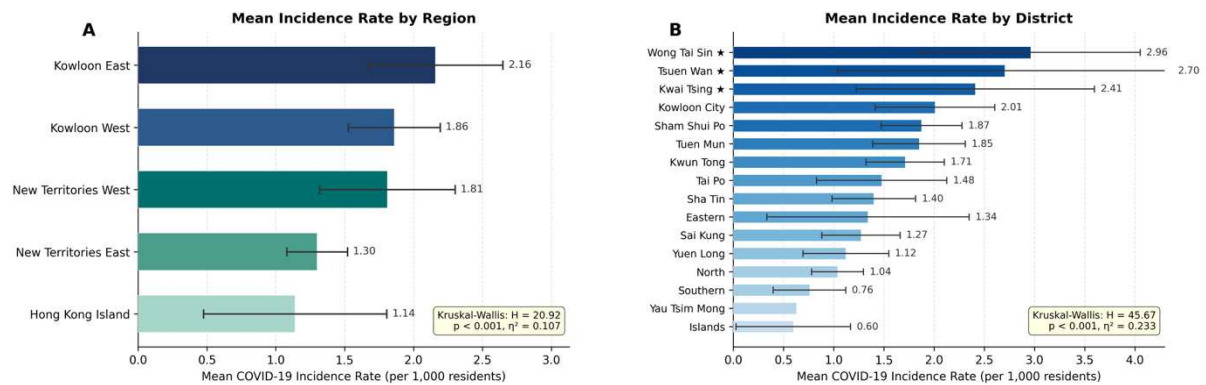


Figure 1. Regional and district-level cumulative COVID-19 incidence across public housing estates in Hong Kong. Mean estate-level COVID-19 incidence rates are shown by (A) region and (B) district, with error bars representing 95% confidence intervals. Stars indicate the three districts with the highest mean incidence rates.

The Kruskal-Wallis test showed significant differences in COVID-19 incidence across regions ($H = 20.92$, $df = 4$, $p < 0.001$, $\eta^2 = 0.107$), as shown in **Figure 1A**, and districts ($H = 45.67$, $df = 15$, $p < 0.001$, $\eta^2 = 0.233$) in **Figure 1B**, indicating, respectively, moderate and large effect sizes (**Table S1-S4**). The Kowloon East region had the highest incidence (2.16 per 1,000), nearly 1.9 times that of the Hong Kong Island region (1.14 per 1,000). Dunn's post hoc test with Bonferroni correction identified three significant pairwise regional differences: Hong Kong Island versus Kowloon East ($p = 0.003$), Hong Kong Island versus Kowloon West ($p = 0.006$), and Kowloon East versus New Territories East ($p = 0.032$). At the district level, Wong Tai Sin had the highest mean incidence (2.96 per 1,000), followed by Tsuen Wan (2.70 per 1,000) and Kwai Tsing (2.41 per 1,000), whereas the Islands District had the lowest incidence (0.60 per 1,000).

Global Moran's I indicated significant positive spatial autocorrelation ($I = 0.194$, $z = 4.77$, $p = 0.002$), rejecting the null hypothesis of spatial randomness. LISA identified 10 statistically

significant High-High (HH) hotspot estates, 25 Low-Low (LL) cold spot estates, 4 Low-High (LH), and 4 High-Low (HL) spatial outliers, while 154 estates were non-significant (**Figure 3**). The HH hotspot cluster was concentrated in Wong Tai Sin district (Tsz Ching, Tsz Hong, Tsz Lok, Tsz Man, Fu Shan, and Fung Tak) and along the adjacent Kwai Tsing-Tsuen Wan corridor (Kwai Shing West, Kwai Luen, Tai Wo Hau, and Cheung Shan). In contrast, the LL cold spot cluster was distributed across geographically peripheral areas, including the Islands, Southern, North, Yuen Long, and Sai Kung districts.

3.2.Socioeconomic correlates of estate-level COVID-19 incidence

Spearman's rank correlations between estate-level predictors and COVID-19 incidence were computed across thematic domains in **Figure 4**. Benjamini-Hochberg false discovery rate correction for all 20 simultaneous tests indicates that no predictor met the prespecified significance threshold of $q < 0.10$. This reflects the inherently multivariable structure of COVID-19 transmission risk, where no single predictor independently accounts for the estate-level burden. Nominally significant associations (unadjusted $p < 0.05$) show directional coherence and theoretically meaningful patterns (**Table S5**).

The strongest nominal associations were observed for demographic and housing affordability. Household size has the strongest bivariate association ($\rho = -0.192$, $p = 0.007$), with larger households having a lower incidence, while the rent-to-income ratio shows the strongest positive association ($\rho = 0.182$, $p = 0.011$). A single-person household ($\rho = 0.173$, p

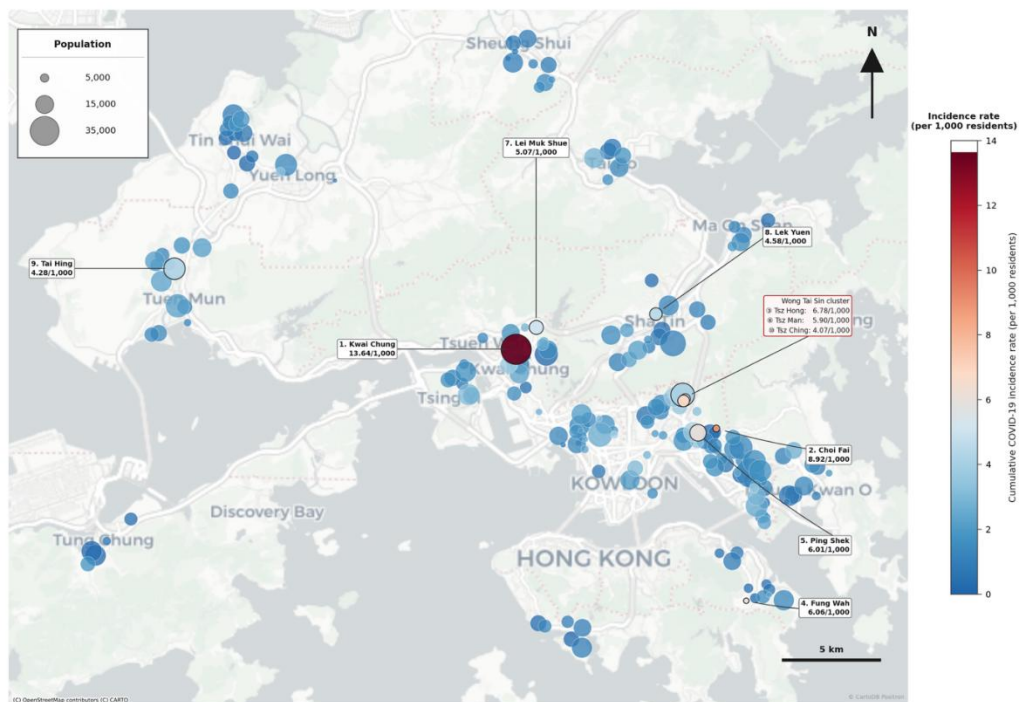


Figure 2. Geospatial distribution of cumulative COVID-19 incidence and population size across 197 public housing estates in Hong Kong. Each bubble represents one public housing estate, with bubble size proportional to estate population and color indicating cumulative COVID-19 incidence per 1,000 residents on a blue-to-red gradient. Estates with the ten highest cumulative incidence rates are labeled to highlight spatial concentrations of elevated disease burden



Figure 3. LISA cluster map of cumulative COVID-19 incidence across Hong Kong public housing estates. Statistically significant High-High clusters indicate incidence hot spots (red, $n = 10$), while Low-Low clusters indicate incidence cold spots (dark blue, $n = 25$). Low-High and High-Low spatial outliers are shown in lighter shades, and estates without statistically significant local spatial autocorrelation are shown in grey ($n = 154$).

= 0.015) and the proportion of craft workers ($\rho = 0.164$, $p = 0.022$) also showed positive trends. In contrast, physical environment variables have negligible and insignificant associations ($|\rho| < 0.09$). The full 20-predictor table is provided in **Table S6**.

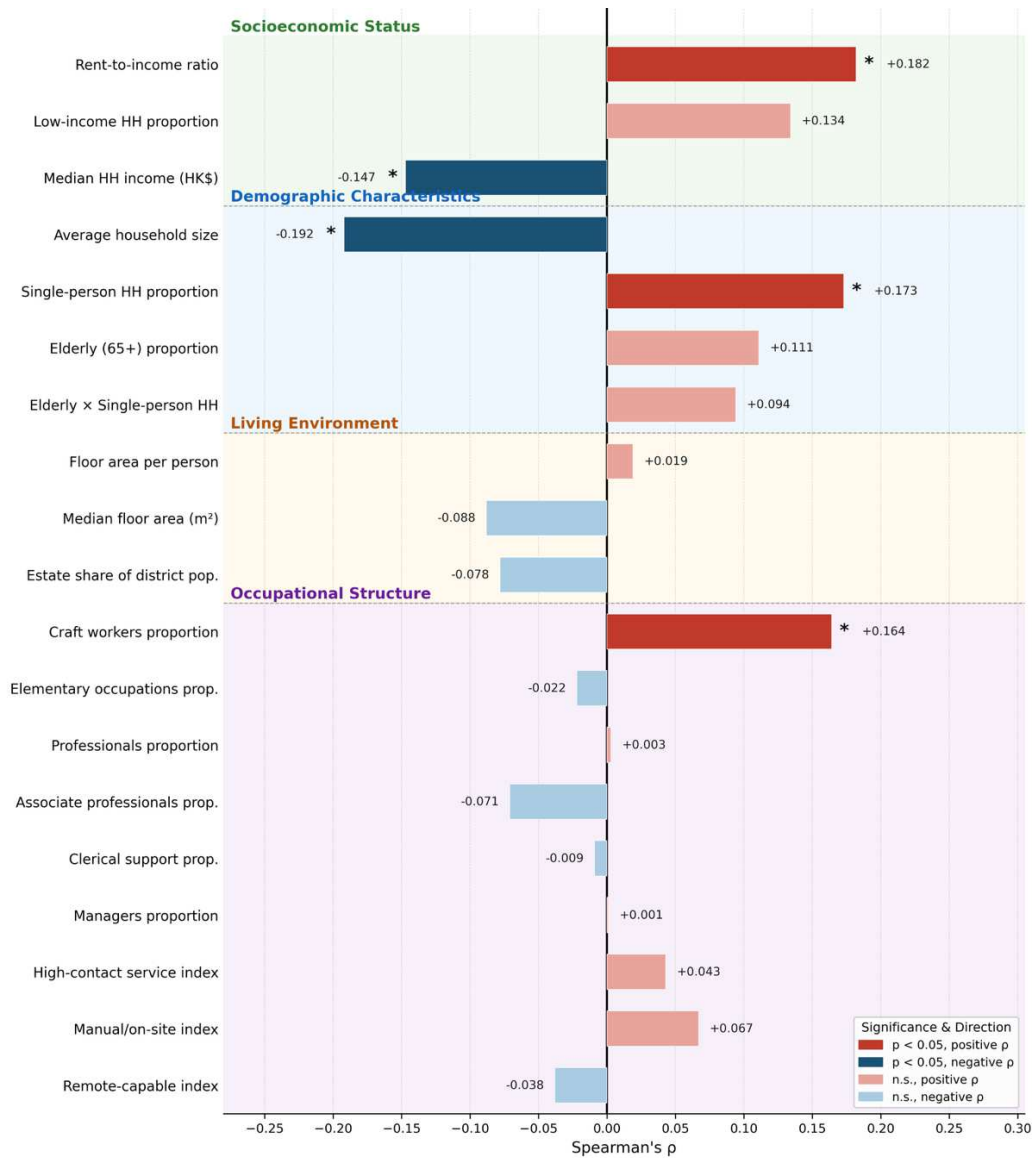


Figure 4. Spearman rank correlations between estate-level socioeconomic predictors and cumulative COVID-19 incidence across 197 public housing estates in Hong Kong, grouped by thematic domain. Horizontal bars represent Spearman's correlation coefficients (ρ), with bar direction indicating the direction of association and bar color denoting nominal statistical significance. Asterisks indicate predictors reaching nominal significance at the unadjusted $p < 0.05$ threshold; no associations remained statistically significant after Benjamini-Hochberg false discovery rate correction.

3.3. Random Forest predictive model and SHAP feature importance

The tuned Random Forest model has an R^2 of 0.225 (RMSE = 0.364; MAE = 0.292), as shown in **Figure 5** and summarized in **Table S7**. The modest R^2 is expected, as unmeasured factors, such as individual mobility, heterogeneity in vaccine uptake, wave-specific variant dynamics, and building ventilation characteristics, are likely to contribute substantially to variation in estate-level incidence beyond structural socioeconomic predictors. Nevertheless, integrated SHAP and LIME were used for in-depth analysis, with SHAP analysis bridging the macro and micro, aggregating population-wide feature impacts, revealing the directionality of socioeconomic features, and confirming the structural drivers of our target metric. Concurrently, LIME case studies elucidate the context-specific reasoning behind localized individual predictions.

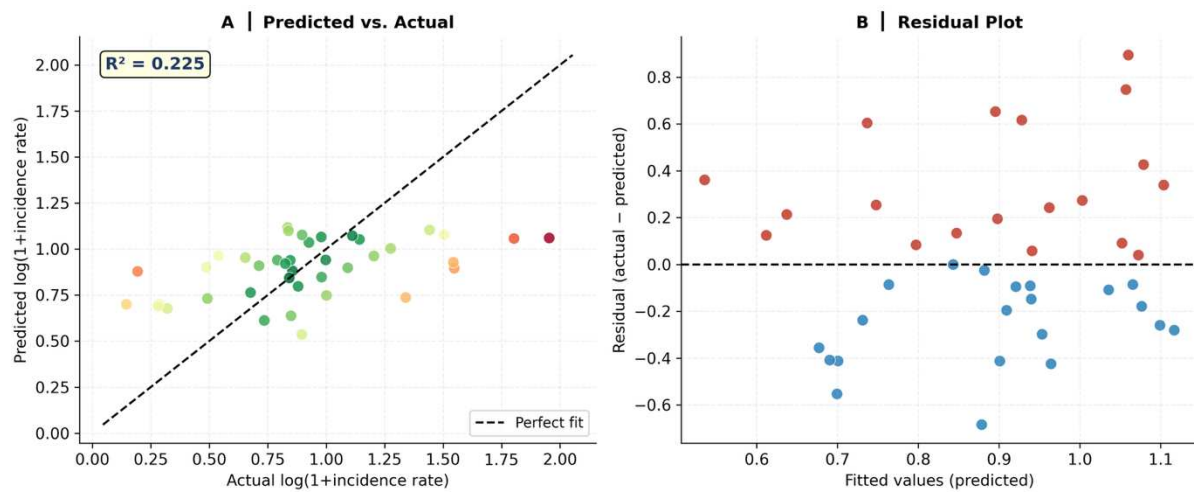


Figure 5. Predictive performance of the Random Forest regression model on the held-out test set. (A) Predicted versus observed $\log(1 + \text{COVID-19 incidence rate})$ across public housing estates. The dashed diagonal line represents perfect agreement between predicted and observed values, with the model explaining 22.5% of the variance in test-set outcomes ($R^2 = 0.225$). (B) Residuals plotted against fitted values. The dashed horizontal line at zero indicates the reference point for prediction errors.

SHAP analysis identified poverty (defined in this study as the proportion of households earning below HK\$10,000 per month) as the most influential predictor of estate-level COVID-19 incidence rate (mean $|\phi| = 0.0445$, 95% CI: 0.040-0.049). Elderly isolation rate ranked second (0.0436; 95% CI: 0.039-0.049), followed by the proportion of craft workers (0.0420;

95% CI: 0.038-0.047). Average household size (0.0377, 95% CI: 0.034-0.041) and rent-to-income ratio (0.0368, 95% CI: 0.033-0.041) formed a second tier of moderately influential predictors (**Figure 6** and **Figure S2**), whereas the remote-capable index (0.0298, 95% CI: 0.027-0.033), overall elderly proportion (0.0260, 95% CI: 0.023-0.029), single-person household share (0.0215, 95% CI: 0.018-0.025), and floor area per person (0.0199, 95% CI: 0.017-0.022) exerted comparatively minor effects on model predictions.

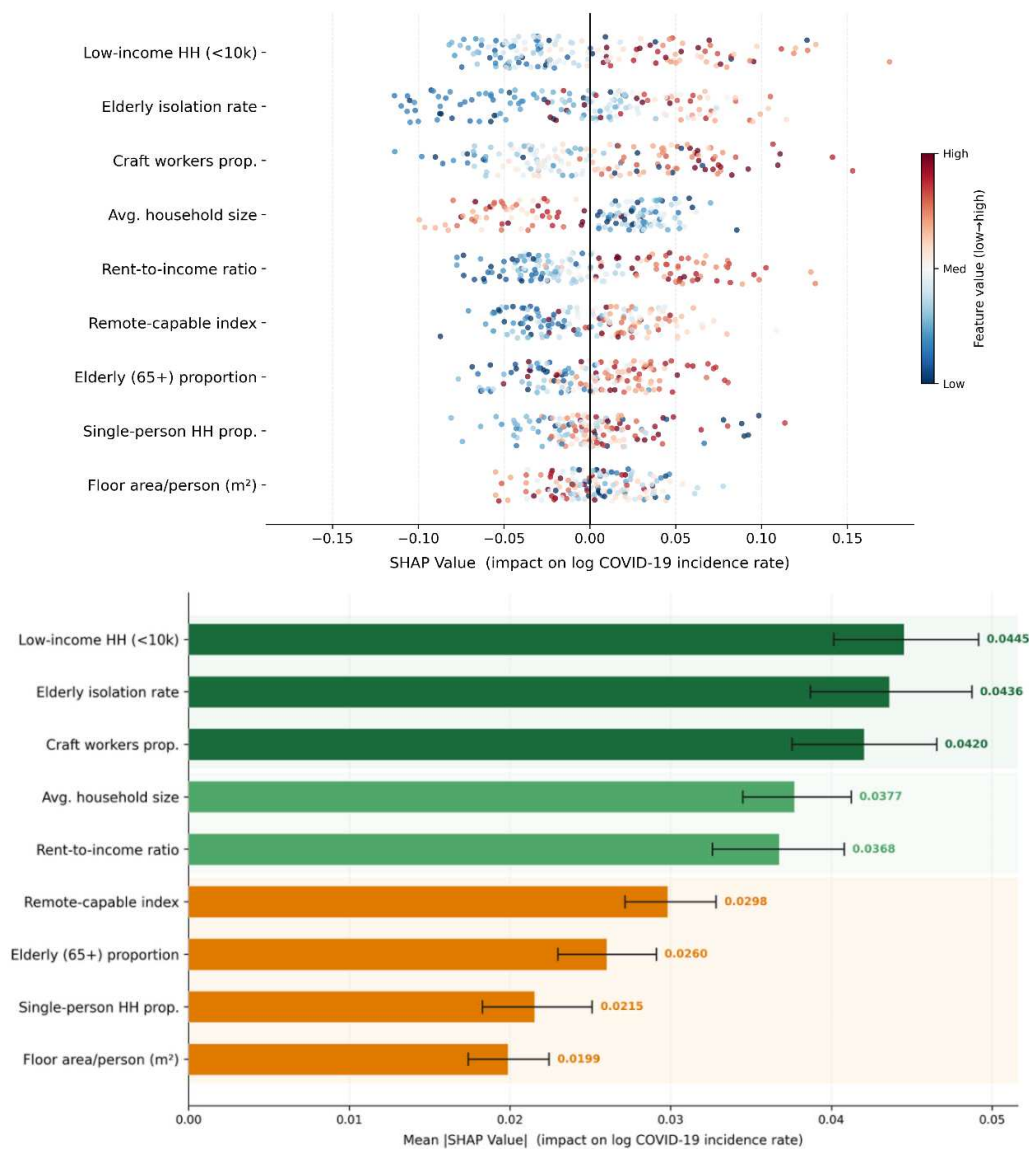


Figure 6. SHAP-based interpretation of the Random Forest model predicting log-transformed COVID-19 incidence across Hong Kong public housing estates. (A) SHAP beeswarm summary plot for the 157 training estates. Each point represents one estate, with the x-axis indicating the SHAP value, corresponding to the feature's contribution to the predicted log incidence rate. Point color denotes the observed feature value, with red indicating higher values and blue indicating lower values. (B) Mean absolute SHAP values for the nine model predictors, with 95% bootstrap confidence intervals, indicating their relative contributions to model predictions.

Three key observations from the SHAP beeswarm plot (**Figure 6A**) demonstrate the model's interpretive structure. First, poverty (rank 1) has a predominantly positive effect: high values (red points) correspond to strongly positive SHAP contributions, indicating that estates with more low-income households consistently have higher COVID-19 incidence, in line with the literature [4,24,25]. Second, the elderly isolation rate (rank 2) and the proportion of craft workers (rank 3) have the same importance, shifting SHAP values upward and contributing to the predicted risk. Comparable risk levels emerge via several substitutable pathways, including economic deprivation, demographic isolation, and occupational exposure, indicating that transmission is multiply realizable and cannot be attributed to a single factor.

Thirdly, the average household size (rank 4) shows a negative SHAP distribution, indicative of being protective, indexing against lower social isolation. The protective role of larger household size and the risk-amplifying role of elderly isolation recur across estates, indicating that social configuration, rather than the age of elderly residents, is more critical. Notably, the confidence intervals for the top three predictors do not overlap with those of features ranked sixth and below (**Figure 6B**), indicating statistically distinguishable ranks of importance.

3.4. Local explainability through LIME estate case studies

While SHAP models behavior across all estates, LIME focuses on the local contribution of each predictor in an individual estate (**Figure 7**). This approach is epidemiologically important, as housing estates with similar COVID-19 incidence may arrive at the same burden through fundamentally different sets of conditions. LIME is instrumental in revealing this heterogeneity that may be obscured in global summaries, thus providing a more complete scientific explanation of how place-specific social and demographic conditions shape estate-level risk.

Six selected estates comprised Tsz Ching Estate (high risk, 4.07 per 1,000) and Chak On Estate (high risk, 3.32 per 1,000); Wah Sum Estate (intermediate risk, 1.84 per 1,000) and Mei

Tin Estate (intermediate risk, 1.39 per 1,000); and Fu Tung Estate (low risk, 0.57 per 1,000) and Tin Ching Estate (low risk, 0.70 per 1,000).

High-risk estates: Tsz Ching Estate and Chak On Estate

Tsz Ching and Chak On estates have high COVID-19 burden during the pandemic. As shown in **Figure 7A**, Tsz Ching Estate's large share of craft and manual workers (10.20% versus territory mean of 8.91%) is the most important risk-amplifying factor (LIME weight: +0.091; 77th percentile), followed by its elderly isolation rate (+0.067; 77th percentile), and single-person households (+0.033; 88th percentile). Income poverty is not observed to contribute to an increased risk (low-income households: -0.036; 27th percentile). Therefore, its high COVID-19 incidence is due to its occupational composition and solo living rather than to poverty.

A plausible exposure mechanism is that non-remote manual employment will have sustained external, physically proximate exposures that were subsequently imported into a community with an unusually high number of socially isolated and single-person households. This pattern suggests that elevated estate-level incidence can arise through the interaction between occupational exposure and socially fragmented household structure, even in the absence of strong poverty-related effects, emphasizing the importance of pathway-specific interventions rather than treating high-risk estates as socially homogeneous.

Figure 7B shows that Chak On Estate's high proportion of low-income households has the largest positive attribution (+0.101, 90th percentile). The estate also had a noticeably smaller average household size of 2.20 people (+0.064; 4th percentile), and a high rate of elderly isolation (+0.062; 88th percentile). In contrast, both craft workers (-0.059; 4th percentile) and remote work (-0.063; 25th percentile) were not significant contributors. Thus, poverty, aging, and social isolation are the main determinants, indicating intra-household and care-related

contact networks as more plausible transmission pathways, particularly for households with elderly residents.

Two sets of different drivers lead to high COVID-19 incidence in these two estates: occupational exposure in Tsz Ching Estate in comparison to poverty and isolation in Chak On Estate. They provide mechanistic insights that aggregate SHAP cannot discern, demonstrating the need for tailored interventions, namely occupational health and infection-control measures for Tsz Ching Estate and elderly welfare outreach and economic support for Chak On Estate.



Figure 7. LIME for six representative public housing estates across the COVID-19 incidence risk spectrum: (A) Tsz Ching Estate and (B) Chak On Estate with high risk levels, (C) Wah Sum Estate and (D) Mei Tin Estate representing estates with moderate risk, (E) Fu Tung Estate and (F) Tin Ching Estate with low risk. Horizontal bars indicate estate-specific local feature weights contributing to the predicted log-transformed incidence rate. Red bars denote predictors that increase the predicted incidence for a given estate, whereas green bars denote predictors that decrease it.

Mid-risk estates: Wah Sum Estate and Mei Tin Estate

Wah Sum Estate had 12.51% of residents employed as craft workers (+0.088; 97th percentile), which is among the highest in the territory, with an additional positive contribution from many remote-capable workers (+0.058; 60th percentile), as shown in **Figure 7C**. The estate had a large household size of 3.0 people (-0.078; 85th percentile) and a low rent burden (-0.047; 11th percentile). Although Wah Sum Estate shared Tsz Ching Estate's occupational-exposure pathway, it remained in the moderate-incidence range due to its larger household size, a lower rent-to-income ratio, and having fewer elderly residents living in isolation. Thus, heightened risks from occupational exposure are insufficient to drive higher COVID-19 burden in the absence of compounding household and socioeconomic vulnerabilities.

Mei Tin Estate had a high proportion (ca. 26.05%) of low-income households (+0.099; 97th percentile), thus a heavier rent burden (rent-to-income ratio = 14.70; +0.099; 85th percentile) as shown in **Figure 7D**. It also had significant craft workers (+0.087; 85th percentile) and smaller household size (+0.072; 45th percentile). These are partially offset by being able to work remotely (-0.074; 24th percentile) and fewer (ca. 1.19%) elderly residents living alone (-0.017; 26th percentile). This case shows that severely disadvantaged communities need not suffer from a high COVID-19 burden when key structural vulnerability is absent, particularly among isolated elderly residents. It clearly indicates that providing social support structures to vulnerable elderly people is crucial in addressing the burden of infectious diseases.

These examples demonstrate the importance of moving beyond single-factor explanations and instead understanding community transmission as arising from the interaction between material disadvantages and socially embedded population vulnerability.

Low-risk estates: Fu Tung Estate and Tin Ching Estate

Fu Tung Estate residents recorded 0.57 per 1,000 (**Figure 7E**). It had an ideal household size of 3.20 people (-0.091; 95th percentile) with a near absence of solitary-living elderly persons (-0.094; 5th percentile). It also has the lowest number of craft workers (0.036; -0.072; 1st percentile) and low deprivation (-0.060; 7th percentile). A small share of single-person households (+0.037; 5th percentile) and the remote-capable index (+0.034; 97th percentile) contributed to slightly increased risks of contracting COVID-19. Tin Ching Estate's very low number of solitary-living elderly residents (-0.097; 18th percentile) was the main reason for its low infection rate (0.70 per 1,000), along with its remote-capable work index (-0.070; 15th percentile) and low rent burden (-0.047; 26th percentile), despite the smaller household size (+0.070; 46th percentile), as can be seen in **Figure 7F**. The patterns suggest that lower risks in these estates were not due to a single protective attribute, but owed to a combination of favorable household composition, minimal social isolation, and fewer socioeconomic disadvantages.

Across these six estates, the local explanations closely reproduced the global risk gradients: risk-amplifying contributions predominated in high-incidence estates, whereas protective contributions predominated in low-incidence estates. More importantly, they indicate that elevated estate-level risk arose through two recurrent but mechanistically distinct pathways: an occupational-exposure pathway characterized by a high number of manual workers with limited capacity for remote work, and a structural-vulnerability pathway, characterized by the co-occurrence of socioeconomic deprivation, small aging households, and socially isolated older residents. This distinction provides a coherent scientific explanation for why estates with broadly comparable disease burden may nonetheless require different intervention strategies.

4. DISCUSSION

4.1. Geographic inequity and spatial clustering

This study presents a highly detailed, granular analysis of disparities in COVID-19 incidence across Hong Kong's public housing, even across institutionally similar housing environments. The 340-fold variation in estate-level incidence, together with significant Kruskal-Wallis effect sizes that are moderate at the regional level ($\eta^2 = 0.107$) and large at the district level ($\eta^2 = 0.233$), confirms that COVID-19 burden was inequitably distributed across the territory. Geographical clustering is evident (**Figures 2-3**), with highest-burden estates located in Kowloon East (particularly Wong Tai Sin District) and along the Kwai Tsing-Tsuen Wan corridor. These areas are loci of socioeconomic deprivation, high-density built environments, and concentrations of essential workers [4,17,26].

The statistically significant positive spatial autocorrelation (Moran's $I = 0.194$, $p = 0.002$) should be highlighted, as it indicates estates with high COVID-19 incidence are geographically clustered together with other similarly high-burden estates. This pattern is consistent with community-level transmission dynamics, shared occupational exposure networks, and common physical infrastructure, such as public transport corridors and wet markets.

For example, the LISA-identified HH hot spot cluster in Wong Tai Sin, involving six geographically adjacent estates that form a contiguous spatial unit, represents a spatially precise and actionable target for prioritized surveillance and resource allocation in future outbreaks. Conversely, the extensive LL cold spot cluster across peripheral districts (Islands, Southern, and North) may reflect lower community connectivity and, potentially, better living space in newer estates. These spatial patterns align with observations from Singapore and South Korea, where geographic proximity to occupational exposure hubs has been shown to amplify transmission risk in dense public housing environments [27].

4.2. Poverty as a primary structural driver

The Random Forest model in **Figure 5** identified the proportion of low-income households (monthly income < HK\$10,000) as the strongest predictor of estate-level COVID-19 incidence (rank 1; mean $|\phi| = 0.0445$). This finding is consistent with the study's theoretical grounding in the Dahlgren-Whitehead model and with extensive international evidence indicating that income poverty is a fundamental structural determinant of vulnerability to COVID-19. Across the United Kingdom, the United States, and other high-income territories, the most socioeconomically deprived communities consistently experienced disproportionately high burdens of infection and mortality [4,28–30]. These inequities likely arose through multiple intersecting mechanisms by which poverty constrains the capacity to adopt protective behaviors, including limited ability to work from home, inability to afford protective equipment, reliance on shared transport, and limited access to healthcare.

Hong Kong's public housing sets the income eligibility limit at approximately HK\$30,000 per month for a family of four, corresponding to almost two-thirds of the city's median household income for such household size, to identify low-income residents experiencing the greatest financial precarity. The beeswarm plot (**Figure 6**) indicates that higher proportions of low-income households (red points) are associated with consistently large positive SHAP contributions, supporting a clear monotonic relationship between income poverty concentration and predicted COVID-19 risk. This interpretation aligns with the Spearman correlation ($\rho = 0.134$, $p = 0.061$).

LIME case studies demonstrated that low-income was the dominant risk-amplifying feature in Chak On Estate (+0.101, the single largest positive attribution among the six estates), where the estate sits in the 90th percentile of deprivation (**Figure 7B**). It is also a major contributing risk factor in Mei Tin Estate (+0.099), as seen in **Figure 7D**. The agreement between SHAP and LIME on income poverty reinforced poverty as a primary structural driver of estate-level

infectious disease risk, consistent with international findings linking area-level disadvantage to elevated incidence and mortality across Chile, Sweden, the United Kingdom, and the United States [31–35].

Income poverty should be understood not merely as a marker, but as an operative mechanism of risk. It limits their ability to acquire essential items, access timely care, and sacrifice income during quarantine isolation. That this factor emerges so clearly within a comparatively homogeneous public-rental housing stock, where dwelling design and tenure are broadly similar, suggests that economic precarity shapes infection risk beyond the physical environment, consistent with a fundamental-cause interpretation of health inequalities [5,28,29,36]. Although important, the Chak On estate (**Figure 7B**) and Mei Tin estate (**Figure 7D**) LIME case studies show income poverty exerts its strongest effects when it co-occurs with demographic vulnerability, rather than acting in isolation.

4.3. Elderly isolation and the structural vulnerability pathway

Elderly residents living alone in social isolation (rank 2; mean $|\phi| = 0.0436$) emerged from LIME analysis as the most pivotal single feature in COVID-19 risk. Crucially, the elderly resident population ranks lower (rank 6; mean $|\phi| = 0.026$), indicating that it is not age but their social isolation that is the source of elevated risk. This distinction is often ignored in conventional bivariate or age-standardized analyses. LIME further identified elderly isolation as a co-dominant amplifier in the high-risk estates (Tsz Ching +0.067 (**Figure 7A**); Chak On +0.063 (**Figure 7B**)) and, in its near absence, acts as a de-risking factor in low-risk estates, Fu Tung Estate (-0.091) and Tin Ching Estate (-0.097), as shown in **Figures 7E-F**.

This observed effect is reasonable, as social isolation aggravates heightened biological susceptibility through immunosenescence and comorbidities [37–39] that require constant access to medicine, nutrition, care, and support [40,41]. These infrequent but often close and

poorly controllable encounters create consequential exposure opportunities leading to a high likelihood of severe COVID-19 infection. Social isolation is a well-established determinant of morbidity and mortality, and population-based cohorts have shown that older adults living alone experienced elevated COVID-19 mortality in high-income settings such as the United Kingdom, Sweden, and Singapore, partly attributable to isolation-related contact seeking and delayed care-seeking [31,42–44].

Our findings extend this evidence to the Hong Kong public-housing context, where the high number of isolated elderly households in older, lower-cost estates may generate a community-level amplification effect that is not detectable in individual-level or bivariate analyses. Taken together, the consistency of the elderly-isolation signal across SHAP and LIME constitutes strong, triangulated evidence that social isolation among older adults represents a primary, independent structural vulnerability. Conceptually, this shifts the focus from age-based risk to relationship-based risk: it is the social embeddedness of older residents, rather than chronological age alone, that shapes estate-level vulnerability.

4.4. Occupational exposure: a clear positive gradient for craft and manual workers

The share of manual and craft workers in an estate (rank 3; mean $|\phi| = 0.0420$) contributes a positive SHAP distribution, aligned with established mechanistic pathways linking non-remote, close-contact manual work to increased community transmission risk [45]. It is a second principal route to elevated estate-level risk, an occupational-exposure pathway that is analytically distinct from the deprivation-isolation axis. It was the single largest risk-amplifying contributor in Tsz Ching Estate (+0.091) and Wah Sum Estate (+0.088), the latter sitting in the 97th percentile of craft-worker share (**Figure 7A** and **7C**). Craft, construction, and manual trades cannot be done remotely. They continued to commute, work in close physical proximity, and congregate when other population groups sheltered in place.

In Fu Tung Estate (**Figure 7E**), which had the territory's lowest craft-worker share (1st percentile), it lowered COVID-19 risks (-0.072). Similarly, in Chak On Estate (**Figure 7B**), the negative attribution (-0.059) correctly indicates that occupation was not the primary driver of heightened risk at this public housing estate. This pattern agrees with the individual-level cohort evidence documenting substantially elevated COVID-19 risk among manual and essential workers [27,45]. It demonstrates that risk arises from how work is organized, where it is performed, and the extent to which it can be relocated away from high-contact settings, rather than from individual choice alone. Future policies must emphasize de-risking essential workers' exposure through workplace-focused measures, strengthened infection prevention and control, improved ventilation and air-quality standards, guaranteed paid sick leave, and income protection that enables symptomatic or exposed workers to temporarily withdraw from work without financial penalties.

The remote-work capability index (ranked 6; mean $|\phi| = 0.0298$) must be understood in an estate-level context. In high-occupational estates, it closely tracked manual-worker exposure, indicating their limited scope for remote work. It measures the extent to which an estate's workforce remained tethered to on-site work with its concomitant exposure risks; it should not imply that remote-work capacity itself is intrinsically hazardous. It covaried with income poverty and occupational structure within a particular housing estate. Where remote work was feasible and adopted at scale, it likely compressed community contact networks and attenuated local transmission, whereas in estates where on-site work and commuting predominated, there were higher exposure risks from workplace and commuting.

4.5. The household size paradox in its proper context

Larger household size (rank 4; mean $|\phi| = 0.0377$) is associated with lower COVID-19 incidence ($\rho = -0.192$, $p = 0.007$). Estates with large household sizes (**Figure 7**), such as Fu Tung Estate (3.20 people) and Wah Sum Estate (3 people), have lower COVID-19 incidence

as compared to Chak On Estate with a smaller household size (2.2 people). This observation appears counterintuitive, as crowding is expected to amplify transmission and contradicts global evidence from countries such as the United Kingdom and the United States, where larger households consistently experienced higher transmission.[46–48]

However, the observed effects of household size must be taken in the context of Hong Kong's specific housing and demographic sociology. Within this segment of the public rental stock, small average household size co-occurs with aging, single-person, and isolated elderly households, and thus operates as a marker of demographic and social vulnerability rather than of spacious living conditions. Larger households, in turn, are predominantly younger, working families in newer estates, who may have greater functional support and fewer underlying frailties despite higher within-household contact rates.

Two pieces of evidence support this interpretation. First, the SHAP dependence plot for household size shows a smooth, monotonic gradient rather than a sharp crowding threshold (**Figure 6**), indicating a graded demographic effect rather than a non-linear density effect. Second, floor area per person, a more direct indicator of physical crowding, was the weakest predictor (rank 9; mean $|\phi| = 0.019$) and did not surface as a major driver of risk. Together, these findings indicate that household size effects emerge mainly from demographic factors, rather than as a proxy for residential crowding, exemplifying a context-specific relationship that would likely be mis-specified by a single-slope linear model but is rendered visible and interpretable through explainable machine learning.

4.6. Limitations

There are several limitations within our interpretations. First, temporal aggregation of cases across epidemiologically distinct waves conflates variation in dominant variants (wild-type to Omicron), vaccination coverage, and the stringency of non-pharmaceutical interventions. Second, the ecological design precludes causal inference at the individual level and is

susceptible to the ecological fallacy. Third, theoretically relevant variables were unavailable at the estate level, including individual mobility patterns, contact network structure, vaccination uptake rates, and building-specific characteristics, which likely account for a portion of unexplained variance. Fourth, LIME's locally linear approximation may inadequately capture non-linearities at feature distribution extremes. Notwithstanding these limitations, the convergence of spatial, correlational, and machine-learning evidence provides strong triangulated support for the principal findings.

5. CONCLUSION

This study demonstrates that COVID-19 incidence across Hong Kong's 197 major public housing estates is geographically clustered and structurally determined by quantifiable socioeconomic inequities operating at the micro-geographic estate level. Global Moran's I confirmed non-random spatial clustering, with LISA identifying a concentrated hot spot in the Wong Tai Sin/Kwai Tsing corridor. The integrated explainable machine learning analysis, combining Random Forest regression with SHAP global importance and LIME local explanations, identifies income poverty, elderly isolation, and craft worker occupational exposure as the principal structural drivers of estate-level COVID-19 incidence. Housing cost burden and the absence of remote-work capacity represent additional modifiable risk dimensions. The model explains 22.5% of estate-level variance, with LIME analysis revealing two mechanistically distinct high-risk pathways: the occupational exposure pathway exemplified by Tsz Ching Estate, where craft and manual worker concentration was the dominant risk-amplifying feature, and a structural vulnerability pathway exhibited by Chak On Estate, where income poverty, small average household size, and elderly isolation jointly dominated.

These findings translate into three evidence-based preparedness priorities: (1) targeted income-support and public health surveillance mechanisms for the most income-deprived estates, particularly within the spatially identified High-High hot spot cluster; (2) proactive elderly isolation outreach programs for estates with high concentrations of elderly single-person households; and (3) workplace infection control and occupational health protections for craft and manual trade workers residing in public housing, whose occupational exposure propagates directly to their residential communities. The integrated geospatial and explainable machine learning framework presented here is reproducible, corrected for methodological limitations, and directly generalizable to comparable high-density public housing.

Author contributions

All authors met the ICMJE criteria for authorship. **JD Pilapil**: Conceptualization, data curation, formal analysis, methodology, writing – original draft, writing – review and editing. **J Woo**: methodology, validation, writing – review and editing. **MP Kwan**: methodology, validation, writing – review and editing. **A Huang**: validation, writing – review and editing. **KL Yeung**: Conceptualization, writing – original draft, writing – review and editing.

Declaration of interest

The authors declare no conflict of interest.

Data availability

COVID-19 surveillance data are publicly available via the Centre for Health Protection of Hong Kong SAR (<https://www.coronavirus.gov.hk>). Socio-demographic profiles of public housing are publicly disseminated by the Hong Kong Housing Authority.

Acknowledgements

This work received crucial financial support from several sources: the Hong Kong Innovation Fund (SST/015/20GP) and the HKUST-AP EnviroSci Ltd Joint Laboratory on Health and Environmental Innovations. We gratefully acknowledge the Centre for Health Protection of Hong Kong SAR for providing the COVID-19 data on which these findings were based.

REFERENCES

1. Commission on Social Determinants of Health. Closing the gap in a generation : health equity through action on the social determinants of health : final report of the commission on social determinants of health [Internet]. Geneva, Switzerland: World Health Organization; 2008 p. 247. <https://iris.who.int/server/api/core/bitstreams/cb08095c-55c8-484e-bff6-0e9c78fd38dd/content>. Accessed 19 May 2026
2. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health [Internet]. Stockholm, Sweden: Institute of Futures Studies; 1991. <https://www.iffs.se/media/1326/20080109110739filmZ8UVQv2wQFShMRF6cuT.pdf>
3. Dahlgren G, Whitehead M. The Dahlgren-Whitehead model of health determinants: 30 years on and still chasing rainbows. Public Health. 2021;199:20–4. <https://doi.org/10.1016/j.puhe.2021.08.009>
4. Robb K, Ahmed R, Wong J, Ladd E, de Jong J. Substandard housing and the risk of COVID-19 infection and disease severity: A retrospective cohort study. SSM - Population Health. 2024;25:101629. <https://doi.org/10.1016/j.ssmph.2024.101629>
5. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health. 2020;74:964–8. <https://doi.org/10.1136/jech-2020-214401>

6. McGowan VJ, Bambra C. COVID-19 mortality and deprivation: pandemic, syndemic, and endemic health inequalities. *The Lancet Public Health*. Elsevier; 2022;7:e966–75. [https://doi.org/10.1016/S2468-2667\(22\)00223-7](https://doi.org/10.1016/S2468-2667(22)00223-7)
7. Patel JA, Nielsen FBH, Badiani AA, Assi S, Unadkat VA, Patel B, et al. Poverty, inequality and COVID-19: the forgotten vulnerable. *Public Health*. 2020;183:110–1. <https://doi.org/10.1016/j.puhe.2020.05.006>
8. Housing Bureau of the HKSAR Government. Housing in Figures 2024 [Internet]. 2024 Nov. <https://www.hb.gov.hk/eng/publications/housing/HIF2024.pdf>
9. Cheung P-HH, Chan ,Chi-Ping, and Jin D-Y. Lessons learned from the fifth wave of COVID-19 in Hong Kong in early 2022. *Emerging Microbes & Infections*. Taylor & Francis; 2022;11:1072–8. <https://doi.org/10.1080/22221751.2022.2060137>
10. Yang B, Lin Y, Xiong W, Liu C, Gao H, Ho F, et al. Comparison of control and transmission of COVID-19 across epidemic waves in Hong Kong: an observational study. *The Lancet Regional Health – Western Pacific* [Internet]. Elsevier; 2024 [cited 2025 Jan 6];43. <https://doi.org/10.1016/j.lanwpc.2023.100969>
11. Fong TCT, Chan CH, Yip PSF. Neighborhood environmental conditions and well-being during the COVID-19 pandemic: A network analysis in Hong Kong adults. *Landscape and Urban Planning*. 2025;254:105247. <https://doi.org/10.1016/j.landurbplan.2024.105247>
12. Qiao Q, Ren C, Chen S, Liang Y, Lai KYY, Zhou Y, et al. Architectural design and building-level infections during the early stage of COVID-19: A study of 2597 public housing buildings in Hong Kong. *Building and Environment*. 2025;276:112853. <https://doi.org/10.1016/j.buildenv.2025.112853>

13. Qiao Q, Sarkar C, Schuldenfrei E, Lai KY, Zhou Y, Webster C. Architectural design and epidemic prevalence: Insights from Hong Kong's fifth wave. *Building and Environment*. 2024;256:111516. <https://doi.org/10.1016/j.buildenv.2024.111516>
14. Saha P, Biswas SK, Biswas MdHA, Ghosh U. An SEQAIHR model to study COVID-19 transmission and optimal control strategies in Hong Kong, 2022. *Nonlinear Dyn*. 2023;111:6873–93. <https://doi.org/10.1007/s11071-022-08181-0>
15. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *The Lancet Public Health*. Elsevier; 2020;5:e279–88. [https://doi.org/10.1016/S2468-2667\(20\)30090-6](https://doi.org/10.1016/S2468-2667(20)30090-6)
16. Yang B, Lin Y, Xiong W, Liu C, Gao H, Ho F, et al. Comparison of control and transmission of COVID-19 across epidemic waves in Hong Kong: an observational study. *The Lancet Regional Health – Western Pacific* [Internet]. Elsevier; 2024 [cited 2025 May 30];43. <https://doi.org/10.1016/j.lanwpc.2023.100969>
17. Chung GK-K, Robinson M, Marmot M, Woo J. Monitoring socioeconomic inequalities in health in Hong Kong: insights and lessons from the UK and Australia. *The Lancet Regional Health – Western Pacific* [Internet]. Elsevier; 2023 [cited 2025 June 1];31. <https://doi.org/10.1016/j.lanwpc.2022.100636>
18. Tan JHI, Goh AXC, Yi H. Health inequity and social determinants of healthcare utilization in the disadvantaged community of Hong Kong. *Health Promotion International*. 2023;38:daab195. <https://doi.org/10.1093/heapro/daab195>
19. Centre for Health Protection. Archives of List of buildings with confirmed / probable cases of COVID-19 [Internet]. Department of Health of the Hong Kong Special Administrative

Region; 2022 [cited 2026 Sept 6]. <https://www.chp.gov.hk/en/features/102991.html>. Accessed 6 Sept 2026

20. Hong Kong Housing Authority. Location and profile of public housing estates [Internet]. DATA.GOV.HK. 2022 [cited 2026 Sept 6]. <https://data.gov.hk/en-data/dataset/hk-housing-eslocator-eslocator>. Accessed 6 Sept 2026

21. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995;57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>

22. Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems* [Internet]. Curran Associates, Inc.; 2017 [cited 2026 May 20]. https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html. Accessed 20 May 2026

23. Ribeiro MT, Singh S, Guestrin C. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* [Internet]. New York, NY, USA: Association for Computing Machinery; 2016 [cited 2026 May 19]. p. 1135–44. <https://doi.org/10.1145/2939672.2939778>

24. Ahmad K, Erqou S, Shah N, Nazir U, Morrison AR, Choudhary G, et al. Association of poor housing conditions with COVID-19 incidence and mortality across US counties. *PLOS ONE*. Public Library of Science; 2020;15:e0241327. <https://doi.org/10.1371/journal.pone.0241327>

25. Wei R, Zhang Y, Gao S, Brown BJ, Hu S, Link BG. Health disparity in the spread of COVID-19: Evidence from social distancing, risk of interactions, and access to testing. *Health & Place*. 2023;82:103031. <https://doi.org/10.1016/j.healthplace.2023.103031>
26. Chen S, Wang S, Sun Y, Liu JYW. Development of a contextualized index of multiple deprivation for age-friendly cities: Evidence from Hong Kong. *Applied Geography*. 2024;167:103285. <https://doi.org/10.1016/j.apgeog.2024.103285>
27. Ng OT, Marimuthu K, Koh V, Pang J, Linn KZ, Sun J, et al. SARS-CoV-2 seroprevalence and transmission risk factors among high-risk close contacts: a retrospective cohort study. *The Lancet Infectious Diseases*. 2021;21:333–43. [https://doi.org/10.1016/S1473-3099\(20\)30833-1](https://doi.org/10.1016/S1473-3099(20)30833-1)
28. Albani V, Welsh CE, Brown H, Matthews FE, Bambra C. Explaining the deprivation gap in COVID-19 mortality rates: A decomposition analysis of geographical inequalities in England. *Social Science & Medicine*. 2022;311:115319. <https://doi.org/10.1016/j.socscimed.2022.115319>
29. Baena-Díez JM, Barroso M, Cordeiro-Coelho SI, Díaz JL, Grau M. Impact of COVID-19 outbreak by income: hitting hardest the most deprived. *J Public Health (Oxf)*. 2020;42:698–703. <https://doi.org/10.1093/pubmed/fdaa136>
30. Lo C-H, Nguyen LH, Drew DA, Warner ET, Joshi AD, Graham MS, et al. Race, ethnicity, community-level socioeconomic factors, and risk of COVID-19 in the United States and the United Kingdom. *eClinicalMedicine* [Internet]. Elsevier; 2021 [cited 2026 May 21];38. <https://doi.org/10.1016/j.eclinm.2021.101029>
31. Drefahl S, Wallace M, Mussino E, Aradhya S, Kolk M, Brandén M, et al. A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. *Nat*

Commun. Nature Publishing Group; 2020;11:5097. <https://doi.org/10.1038/s41467-020-18926-3>

32. Hawkins RB, Charles EJ, Mehaffey JH. Socio-economic status and COVID-19–related cases and fatalities. Public Health. 2020;189:129–34. <https://doi.org/10.1016/j.puhe.2020.09.016>

33. Chen JT, Krieger N. Revealing the Unequal Burden of COVID-19 by Income, Race/Ethnicity, and Household Crowding: US County Versus Zip Code Analyses. Journal of Public Health Management and Practice. 2021;27:S43–56. <https://doi.org/10.1097/PHH.0000000000001263>

34. Karmakar M, Lantz PM, Tipirneni R. Association of Social and Demographic Factors With COVID-19 Incidence and Death Rates in the US. JAMA Netw Open. 2021;4:e2036462. <https://doi.org/10.1001/jamanetworkopen.2020.36462>

35. Mena GE, Martinez PP, Mahmud AS, Marquet PA, Buckee CO, Santillana M. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. Science. American Association for the Advancement of Science; 2021;372:eabg5298. <https://doi.org/10.1126/science.abg5298>

36. Sajjadi S, Toranj Simin P, Shadmangohar M, Taraktas B, Bayram U, Ruiz-Blondet MV, et al. Structural inequalities exacerbate infection disparities. Sci Rep. Nature Publishing Group; 2025;15:9082. <https://doi.org/10.1038/s41598-025-91008-w>

37. Cox LS, Bellantuono I, Lord JM, Sapey E, Mannick JB, Partridge L, et al. Tackling immunosenescence to improve COVID-19 outcomes and vaccine response in older adults. The Lancet Healthy Longevity. Elsevier; 2020;1:e55–7. [https://doi.org/10.1016/S2666-7568\(20\)30011-8](https://doi.org/10.1016/S2666-7568(20)30011-8)

38. Perrotta F, Corbi G, Mazzeo G, Boccia M, Aronne L, D'Agnano V, et al. COVID-19 and the elderly: insights into pathogenesis and clinical decision-making. *Aging Clin Exp Res*. 2020;32:1599–608. <https://doi.org/10.1007/s40520-020-01631-y>
39. Bartleson JM, Radenkovic D, Covarrubias AJ, Furman D, Winer DA, Verdin E. SARS-CoV-2, COVID-19 and the aging immune system. *Nat Aging*. Nature Publishing Group; 2021;1:769–82. <https://doi.org/10.1038/s43587-021-00114-7>
40. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. *The Lancet Public Health*. Elsevier; 2020;5:e256. [https://doi.org/10.1016/S2468-2667\(20\)30061-X](https://doi.org/10.1016/S2468-2667(20)30061-X)
41. Lebrasseur A, Fortin-Bédard N, Lettre J, Raymond E, Bussi eres E-L, Lapierre N, et al. Impact of the COVID-19 Pandemic on Older Adults: Rapid Review. *JMIR Aging*. JMIR Publications Inc., Toronto, Canada; 2021;4:e26474. <https://doi.org/10.2196/26474>
42. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review. *Perspect Psychol Sci*. SAGE Publications Inc; 2015;10:227–37. <https://doi.org/10.1177/1745691614568352>
43. Yip W, Ge L, Ho AHY, Heng BH, Tan WS. Building community resilience beyond COVID-19: The Singapore way. *The Lancet Regional Health – Western Pacific* [Internet]. Elsevier; 2021 [cited 2026 May 21];7. <https://doi.org/10.1016/j.lanwpc.2020.100091>
44. Wing K, Grint DJ, Mathur R, Gibbs HP, Hickman G, Nightingale E, et al. Association between household composition and severe COVID-19 outcomes in older people by ethnicity: an observational cohort study using the OpenSAFELY platform. *Int J Epidemiol*. 2022;51:1745–60. <https://doi.org/10.1093/ije/dyac158>

45. Mutambudzi M, Niedzwiedz CL, Macdonald EB, Leyland AH, Mair FS, Anderson JJ, et al. Occupation and risk of severe COVID-19: prospective cohort study of 120,075 UK Biobank participants. *Occup Environ Med*. 2021;78:307–14. <https://doi.org/10.1136/oemed-2020-106731>
46. Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. Household Transmission of SARS-CoV-2: A Systematic Review and Meta-analysis. *JAMA Netw Open*. 2020;3:e2031756. <https://doi.org/10.1001/jamanetworkopen.2020.31756>
47. Lopez Bernal J, Panagiotopoulos N, Byers C, Garcia Vilaplana T, Boddington N, Zhang X-S, et al. Transmission dynamics of COVID-19 in household and community settings in the United Kingdom, January to March 2020. *Eurosurveillance* [Internet]. 2022 [cited 2025 June 1];27. <https://doi.org/10.2807/1560-7917.ES.2022.27.15.2001551>
48. Grijalva CG, Rolfes MA, Zhu Y, McLean HQ, Hanson KE, Belongia EA, et al. Transmission of SARS-COV-2 Infections in Households — Tennessee and Wisconsin, April–September 2020. *MMWR Morb Mortal Wkly Rep* [Internet]. 2020 [cited 2025 June 1];69. <https://doi.org/10.15585/mmwr.mm6944e1>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)