Relationship of various COVID-19 antibody titer with individual characteristics and prediction of future epidemic trend in Xiamen City, China

Liansheng Xu[1#], Buasiyamu Abudunaibi[2#], Zhiqi Zeng[3#], Yunkang Zhao[2], Yao Wang[2], Xiaohao Guo[2], Yidun Zhang[1], Tao Li[2], Wenkui Lu[1], Weiliang Tian[4], Zhinan Guo[1*], Chenghao Su[5*], Tianmu Chen[2*]

[1]Xiamen Center for Disease Control and Prevention, Xiamen, China; [2]State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, School of Public Health, Xiamen University, Xiamen, China; [3]Respiratory Disease AI Laboratory on Epidemic Intelligence and Medical Big Data Instrument Applications, Macao University of Science and Technology, Macao, China; [4]School of Medicine, Duke University, Durham, NC, USA; [5]Zhongshan Hospital, Fudan University (Xiamen Branch), Xiamen, China

Contributions: (I) Conception and design: T Chen, L Xu, B Abudunaibi, Y Zhao, Y Wang; (II) Administrative support: T Chen, L Xu, Z Guo, C Su, Z Zeng; (III) Provision of study materials or patients: Z Guo, C Su, L Xu, Y Zhang, W Lu; (IV) Collection and assembly of data: L Xu, Y Zhang, W Lu; (V) Data analysis and interpretation: B Abudunaibi, Y Zhao, T Li, Y Wang, X Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

[*]These authors contributed equally to this work as co-senior authors.

Correspondence to: Zhinan Guo, MD. Xiamen Center for Disease Control and Prevention, 685, Shengguang, Jimei District, Xiamen 361000, China. Email: guozhinan@hotmail.com; Chenghao Su, MD. Zhongshan Hospital, Fudan University (Xiamen Branch), 668, Jinhu Road, Huli District, Xiamen 361000, China. Email: 1272208372@qq.com; Tianmu Chen, PhD. State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, School of Public Health, Xiamen University, 4221-117 South Xiang'an Road, Xiang'an District, Xiamen 361000, China. Email: 13698665@qq.com; chentianmu@xmu.edu.cn.

Background: Reinfection of coronavirus disease 2019 (COVID-19) has raised concerns about how reliable immunity from infection and vaccination is. With mass testing for the virus halted, understanding the current prevalence of COVID-19 is crucial. This study investigated 1,191 public health workers at the Xiamen Center for Disease Control, focusing on changes in antibody titers and their relationship with individual characteristics.

Methods: The study began by describing the epidemiological characteristics of the study participants. Multilinear regression (MLR) models were employed to explore the associations between individual attributes and antibody titers. Additionally, group-based trajectory models (GBTMs) were utilized to identify trajectories in antibody titer changes. To predict and simulate future epidemic trends and examine the correlation of antibody decay with epidemics, a high-dimensional transmission dynamics model was constructed.

Results: Analysis of epidemiological characteristics revealed significant differences in vaccination status between infected and non-infected groups (χ^2=376.706, P<0.05). However, the distribution of antibody titers among the infected and vaccinated populations was not significantly different. The MLR model identified age as a common factor affecting titers of immunoglobulin G (IgG), immunoglobulin M (IgM), and neutralizing antibody (NAb), while other factors showed varying impacts. History of pulmonary disease and hospitalization influenced IgG titer, and factors such as gender, smoking, family history of pulmonary diseases, and hospitalization impacted NAb titers. Age was the sole determinant of IgM titers in this study. GBTM analysis indicated a "gradual decline type" trajectory for IgG (95.65%), while IgM and NAb titers remained stable over the study period. The high-dimensional transmission dynamics model predicted and simulated peak epidemic periods in Xiamen City, which correlated with IgG decay. Age-group-specific simulations revealed a higher incidence and infection rate among individuals aged 30–39 years during both the second and third peaks, followed by those aged 40–49, 50–59, 18–29, and 70–79 years.

Conclusions: Our study shows that antibody titer could be influenced by age, previous pulmonary diseases as well as smoking. Furthermore, the decline in IgG titers is consistent with epidemic trends. These findings emphasize the need for further exploration of these factors and the development of optimized self-protection countermeasures against reinfection.

Keywords: Coronavirus disease 2019 (COVID-19); antibody titer; reinfection; statistical models; transmission dynamics model

Introduction

The global impact of coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has been profound, and efforts to mitigate its effects have been ongoing for over 3 years (1). As of 13 August 2023, over 769 million confirmed cases and over 6.9 million deaths have been reported globally (1). Since December 2022, most countries have been adapting their emergency responses, and attention has turned to understanding the duration of the immune response to infection (2). As vaccination remains a cornerstone for controlling the COVID-19 infection, antibodies obtained from infection also play a role in warding off reinfection (3).

The humoral immune responses, whether triggered by

COVID-19 infection or vaccination, are crucial to fight a second or third infection in the current massive reinfection circumstances (4). Humoral immunity is characterized by the production of antibodies by B cells as a response to antigens. Researchers concluded that after stimulating the immune reaction to COVID-19, the first antibody to appear is immunoglobulin M (IgM), succeeded by immunoglobulin A (IgA) and, notably, immunoglobulin G (IgG). However, the duration of these antibody types varies (5). IgM, for example, typically persists for only 20–30 days, while IgG, detectable approximately 10–14 days post-infection and peaking around day 25, provides longer-lasting immunity. Serum IgA levels rapidly decline, with seropositivity diminishing after 2 months (4), although neutralizing IgA can remain detectable in saliva for an extended period (49–73 days post-symptoms) (6). Importantly, a study suggested that IgA is a significant isotype in nearly neutralizing serum activity, whereas IgG is a key component of serum neutralizing antibodies (NAbs) following SARS-CoV-2 infection (7).

NAb titers after mild SARS-CoV-2 infection were reported to be comparable to those induced by first-generation COVID-19 vaccines (8), and second-vaccine doses further enhance cross-neutralizing activity against variants of concern. The persistence of protective immune responses is linked to the time since the initial infection and the robustness of the peak antibody response (2,3,8,9).

With SARS-CoV-2 persisting in the human population for over 3 years, reports of reinfection cases have become increasingly prevalent. Studies about the relationship between antibody duration (arising from infection or vaccination) and reinfection risk have been conducted. A study that followed 12,541 healthcare workers at Oxford University Hospital revealed that the incidence of SARS-CoV-2 was inversely proportional to baseline titers of

Highlight box

Key findings

- A study of 1,191 public health workers unveiled age as a critical factor influencing immunoglobulin G (IgG), immunoglobulin M, and neutralizing antibody (NAb) titers, while past pulmonary conditions and smoking affected specific antibody levels. Notably, IgG decline mirrored epidemic trends, suggesting a potential correlation.

What is known and what is new?

- Factors, such as age, gender, smoking, severity of symptoms and so on, potentially affect antibody level.
- This manuscript uniquely ties IgG decay to epidemic trends, shedding light on its implications for coronavirus disease 2019 dynamics.

What is the implication, and what should change now?

- Understanding age-linked antibody variations and their connection to infection trends calls for tailored strategies against reinfection. Further exploration of influential factors and targeted protective measures are imperative in shaping effective long-term defenses.

antibodies (10). Another study about the reinfection of SARS-CoV-2 implicated that herd immunity from infection is unlikely to be sufficient to eliminate the virus if reinfections are common (11). However, it is unneglectable that there is no evidence confirming that current intramuscular vaccines can reduce shedding in the event of reinfection. What's more, as SARS-CoV-2 continues to mutate, the immune escape of variants should also be taken into consideration when exploring antibody dynamics (12).

During the COVID-19 pandemic, researchers have come up with various methods to simulate and predict various epidemic scenarios. Transmission dynamics models have played a crucial role in this context (13-15). In light of the current situation, where mass virus testing has been postponed among the general population, constructing models that accurately reflect the real-world situation has become essential.

In this study, we examined the antibody titers of 1,191 public health workers from the Xiamen Center for Disease Control. Our objectives were to investigate potential relationships between individual characteristics and antibody titers, predict future outbreaks by optimizing existing COVID-19 transmission dynamics models, and explore the interplay between epidemic peaks and the decay of specific antibody types within the population. We present this article in accordance with the TRIPOD reporting checklist (available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-1516/rc).

Methods

Study design

This study aims to investigate antibody titer dynamics among Xiamen healthcare workers through three key stages. First, we collected data via questionnaires distributed among healthcare staff in the Xiamen Healthcare System. This yielded 1,191 eligible participants, selected based on SARS-CoV-2 infection or vaccination status post-December 2022. Subsequently, serum antibody testing occurred at monthly intervals from February to May to track titer variations. Finally, we employed a high-dimensional transmission dynamics model to predict future epidemic trends, examining correlations between antibody decay and epidemic peaks. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of the Xiamen

Center for Disease Control [XJK/LLSC(2022)004] and participants signed the agreement consent before filling out the questionnaire.

Data collection

This study collected data in two stages. Initially, questionnaires were distributed among healthcare workers in Xiamen City, gathering individual information, health history, and COVID-19 vaccination records. For those previously infected with SARS-CoV-2, additional infection details were gathered. Based on prior research indicating that NAbs for COVID-19 might persist for a maximum of 73 days (6), questionnaires that were incomplete or lacking essential infection-related information were excluded from the study, and 1,191 valid samples remained. Subsequently, a monthly serum antibody test tracked antibody decay from February to May.

Explore the epidemiological characteristics as well as correlation of individual properties and antibody titer for COVID-19

This study investigated the correlation between individual baseline properties (age, height, weight, and previous health conditions) and COVID-19 antibody titers. Multivariate regression models were established to analyze the relationship, with regression coefficients (β), 95% confidence intervals (CIs), and standardized regression coefficients (β) calculated. Before establishing the regression model, we carried out variable distribution description and Spearman's correlation analysis, and eliminated variables with strong correlation (r>0.7). The optimal model was selected using backward elimination, Bayesian information criterion (BIC), and Akaike information criterion (AIC).

The formula of the model is as follows:

$$\begin{aligned} Y_{IgG} = f_0 &+ S(sex) + f(age) + S(loca) + f(height) \\ &+ f(weight) + S(smk) + S(resd) + S(cd) \\ &+ S(vac) + S(phsms) + S(PI) + S(symp) \\ &+ S(hosp) + f(SI) + S(his) \end{aligned} \quad [1]$$

$$\begin{aligned} Y_{IgM} = f_0 &+ S(sex) + f(age) + S(loca) + f(height) \\ &+ f(weight) + S(smk) + S(resd) + S(cd) \\ &+ S(vac) + S(phsms) + S(PI) + S(symp) \\ &+ S(hosp) + f(SI) + S(his) \end{aligned} \quad [2]$$
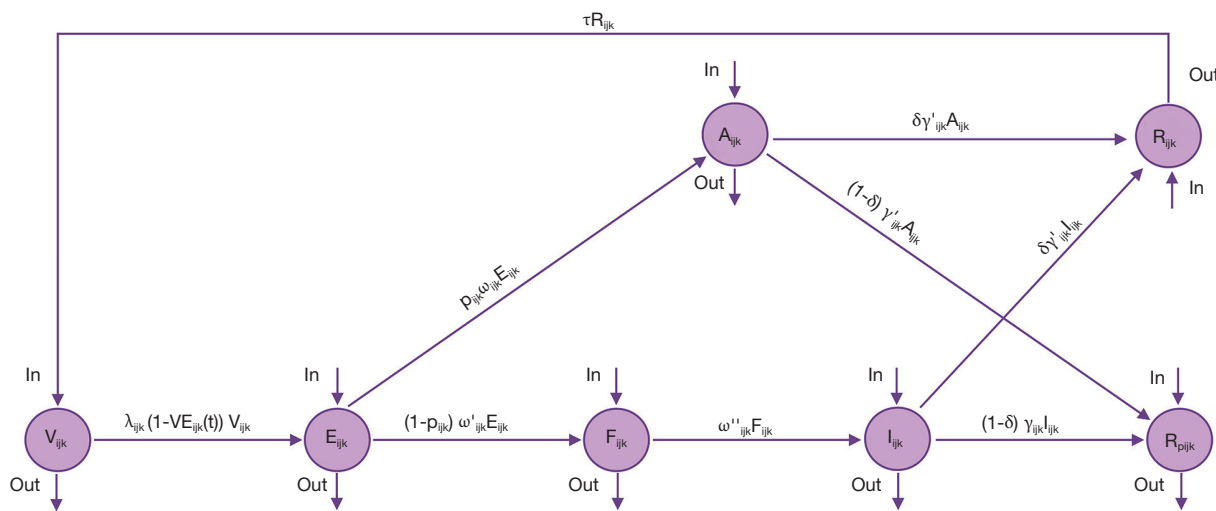
Figure 1 The framework for the VEAFIRPRV model. Arrows among compartments show the flow of the population, while arrows marked "In" and "Out" represent the natural demographic changes in the population. "In" includes immigration as well as newborns, while "Out" includes migrating out and natural death.

$$Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) \\ + f(weight) + S(smk) + S(resd) + S(cd) \\ + S(vac) + S(phsms) + S(PI) + S(symp) \\ + S(hosp) + f(SI) + S(his) \qquad [3]$$

A group-based trajectory model (GBTM) was applied to analyze the longitudinal serum antibody data (16). The GBTM model identified potential clusters with similar antibody trajectory patterns. The study hypothesized that the population's serum could be divided into up to five categories, namely, gradual growth type, gradual decline type, unchanged type, growth and then decline type, and decline and then growth type, reflecting different patterns of change in antibody levels. The highest order of the model was set to 3, i.e., the number of potential subgroups in the population was 1–5 groups, 0–3 orders, with the order reflecting the degree of rapidity or slowness of the trend of change. BIC and AIC were used to assess the model's fit, with an emphasis on minimizing the absolute value of the index.

χ^2 analysis and ANOVA were used to compare demographic differences between different serologic change patterns. Multiple linear and logistic regression analyses were conducted to further explore demographic differences. Statistical significance was set at $P < 0.05$.

Prediction and simulation of the current COVID-19 epidemic by constructing a high-dimensional transmission dynamics model

According to the natural history and vaccination situation for COVID-19 in Xiamen City, we have constructed a high-dimensional transmission dynamics model. This is a well-organized ordinary differential equation (ODE) model in tensor form, which includes factors of age, vaccine, regional contact, stages in disease progress, and critical events such as population growth, and migration. In this model, we grouped the total population N into susceptible, V; exposed, E; asymptomatic infection, A; pre-symptomatic infection, F; symptomatic infection, I; removed/recovered that will not be reinfected, R_p; recovered and will possibly be reinfected, R; therefore, this high-dimensional transmission dynamics model could also be named as VEAFIR$_p$RV model (*Figure 1*).

The model settings are detailed in Appendix 1.

All the values of parameters in the model are presented in *Table 1*.

The functions for the high-dimensional transmission dynamics model (VEAFIR$_p$RV model) are as follows:

$$\frac{d}{dt}V_{ijk} = \tau R_{ijk} - \lambda_{ijk}\left(1 - VE_{ijk}(t)\right)V_{ijk} \qquad [4]$$

Table 1 Definition and values of parameters in VEFIARpRV model of COVID-19

Parameters	Definition	Unit	Value	Source
p	Proportion of asymptomatic infections	1	0.08	Reference (17-21)
ω	Inverse of the average latent period	Day-1	1/3	
ω'	Inverse of the average incubation period from exposed (e) to pre-symptomatic infection (F)	Day-1	1/2	
ω"	Inverse of the average incubation period from pre-symptomatic infection (F) to infections (I)	Day-1	1/3	
γ	Recovery or removal rate for symptomatic infections	1	1/6.8	
γ'	Recovery or removal rate for asymptomatic infections	1	1/6.8	
τ	Inverse of the average duration of losing immunity	Day-1	1/180	
δ	Proportion of those become susceptible again after recovery	1	0.7	

COVID-19, coronavirus disease 2019.

$$\frac{d}{dt}E_{ijk} = \lambda_{ijk}\left(1-VE_{ijk}(t)\right)V_{ijk} - p_{ijk}\omega_{ijk}E_{ijk} - \left(1-p_{ijk}\right)\omega'_{ijk}E_{ijk} \quad [5]$$

$$\frac{d}{dt}A_{ijk} = p_{ijk}\omega_{ijk}E_{ijk} - \left(1-\delta\right)\gamma'_{ijk}A_{ijk} - \delta\gamma'_{ijk}A_{ijk} \quad [6]$$

$$\frac{d}{dt}F_{ijk} = \left(1-p_{ijk}\right)\omega'_{ijk}E_{ijk} - \omega''_{ijk}F_{ijk} \quad [7]$$

$$\frac{d}{dt}I_{ijk} = \omega''_{ijk}F_{ijk} - \left(1-\delta\right)\gamma_{ijk}I_{ijk} - \delta\gamma_{ijk}I_{ijk} \quad [8]$$

$$\frac{d}{dt}R_{p_{ijk}} = \left(1-\delta\right)\gamma_{ijk}I_{ijk} + \left(1-\delta\right)\gamma'_{ijk}A_{ijk} \quad [9]$$

$$\frac{d}{dt}R_{ijk} = \delta\gamma_{ijk}I_{ijk} + \delta\gamma'_{ijk}A_{ijk} - \tau R_{ijk} \quad [10]$$

After building the model, we applied it in the prediction and simulation for the future COVID-19 epidemic. We used the incidence rate and infection rate as the indexes for the scale of the COVID-19 epidemic in Xiamen City and carried out interventions that meet the current policies.

Statistical analysis

Data entry and organization related to this study were performed in Excel 2019. Continuous quantitative variables were described by median ± interquartile range (IQR), and categorical qualitative variables by percentages. Statistical analysis was performed by SPSS version 22.0, and differences were statistically significant at $P<0.05$. For a cross-sectional study, a multivariant regression model was constructed via *Python's statsmodels* package. Then longitudinal data was analyzed by applying the GBTM model on *traj* package of Stats 17.0. High-dimensional

transmission model was processed on Matlab R2023a, differential equations were solved using the fourth-order Runge Kutta method, and model convergence was based on the least root mean square (LRMS), further using the coefficient of determination (R^2) to determine the goodness of fit. Graphs were plotted using CorelDRAW 2020.
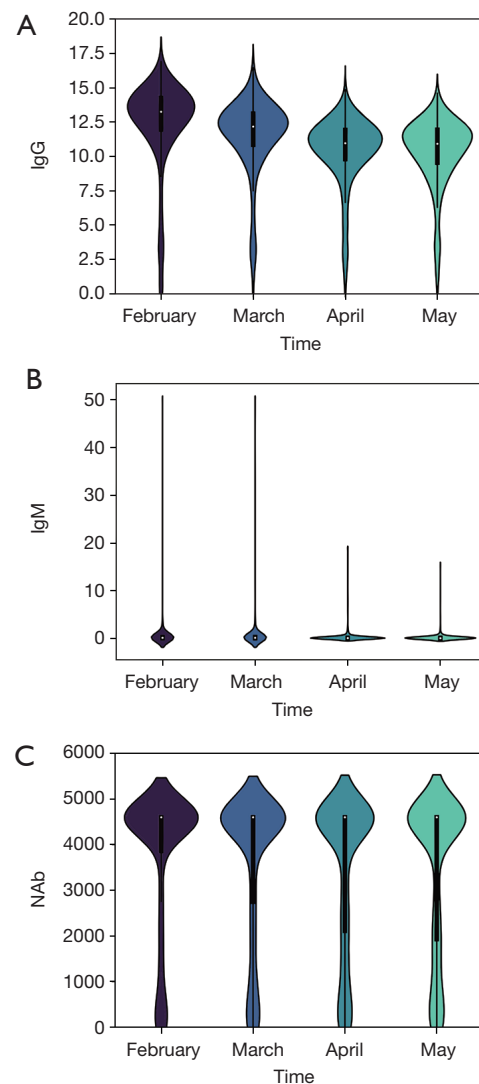
Results

Crowd distribution for the samples

On February 10th, 2023, we initiated data collection by distributing questionnaires to healthcare workers in Xiamen City, including staff from hospitals, Centers for Disease Control and Prevention (CDCs), and primary healthcare organizations. Since it is difficult to collect both infection and serum antibody information from the general population or implement random sampling in the targeted healthcare workers, we have implemented convenience sampling. These questionnaires aimed to collect baseline information about the participants, which included details about their health conditions, COVID-19 vaccination status, and information regarding their previous SARS-CoV-2 infections, such as symptoms and infection dates. Subsequently, serum antibody tests were conducted on the participants. A total of 1,344 questionnaires were collected along with their corresponding serum antibody results. To establish clear inclusion and exclusion criteria for our study, we considered that serum antibodies could persist for a maximum of 73 days (6). We included individuals who had developed antibodies within 73 days before serum

Table 2 Crowd distribution for the samples

Categories	Non-infected	Infected	Tests
Total, n	135	1,056	
Gender, n			χ^2=0.09, P=0.92
Male	39	301	
Female	96	755	
Age (years), n			χ^2=9.639, P=0.08
<20	0	3	
21–30	48	263	
31–40	50	462	
41–50	23	233	
51–60	14	87	
>61	0	6	
Address, n			χ^2=30.053, P<0.05
Jinjiang City	0	1	
Haicang	8	79	
Huli	21	118	
Jimei	32	140	
Siming	36	517	
Tongan	22	105	
Xiangan	16	95	
Longhai	0	1	
Vaccination, n			χ^2=376.706, P<0.05
Fundamental immune	1	61	
Booster 1	7	676	
Booster 2	87	304	
Booster 3	40	7	

Figure 2 Distribution of antibody titers acquired from four serum tests. (A) The figure of IgG (COI), differences in color and X-axis both represent different months that the test was conducted, which was the same in the other two types of antibodies. For (B) IgM (COI) and (C) NAb (IU/mL). IgG, immunoglobulin G; IgM, immunoglobulin M; NAb, neutralizing antibody; COI, cut off index.

testing. For those who were infected by SARS-CoV-2, we considered the day of symptom onset as the antibody-emerging date, while for those who were not infected, we used the last vaccination date as the antibody-emerging date. As a result, 1,191 participants met our inclusion criteria and were included in the study.

Of the 1,191 participants included in the study, 1,056 individuals were confirmed to be infected with COVID-19 (infections), and 135 individuals were not infected (no infections), as indicated in *Table 2*. Statistical analysis revealed that there were no significant differences in terms of gender (χ^2=0.09, P=0.92) and age groups (χ^2=9.639,

P=0.08) between infected and non-infected individuals. However, there is statistical significance in the difference of the addresses (χ^2=30.053, P<0.05) and vaccination status between infected and non-infected groups (χ^2=376.706, P<0.05). Notably, the distributions of antibody titers for IgG, IgM, and NAb were found to be similar, as depicted in *Figure 2*.
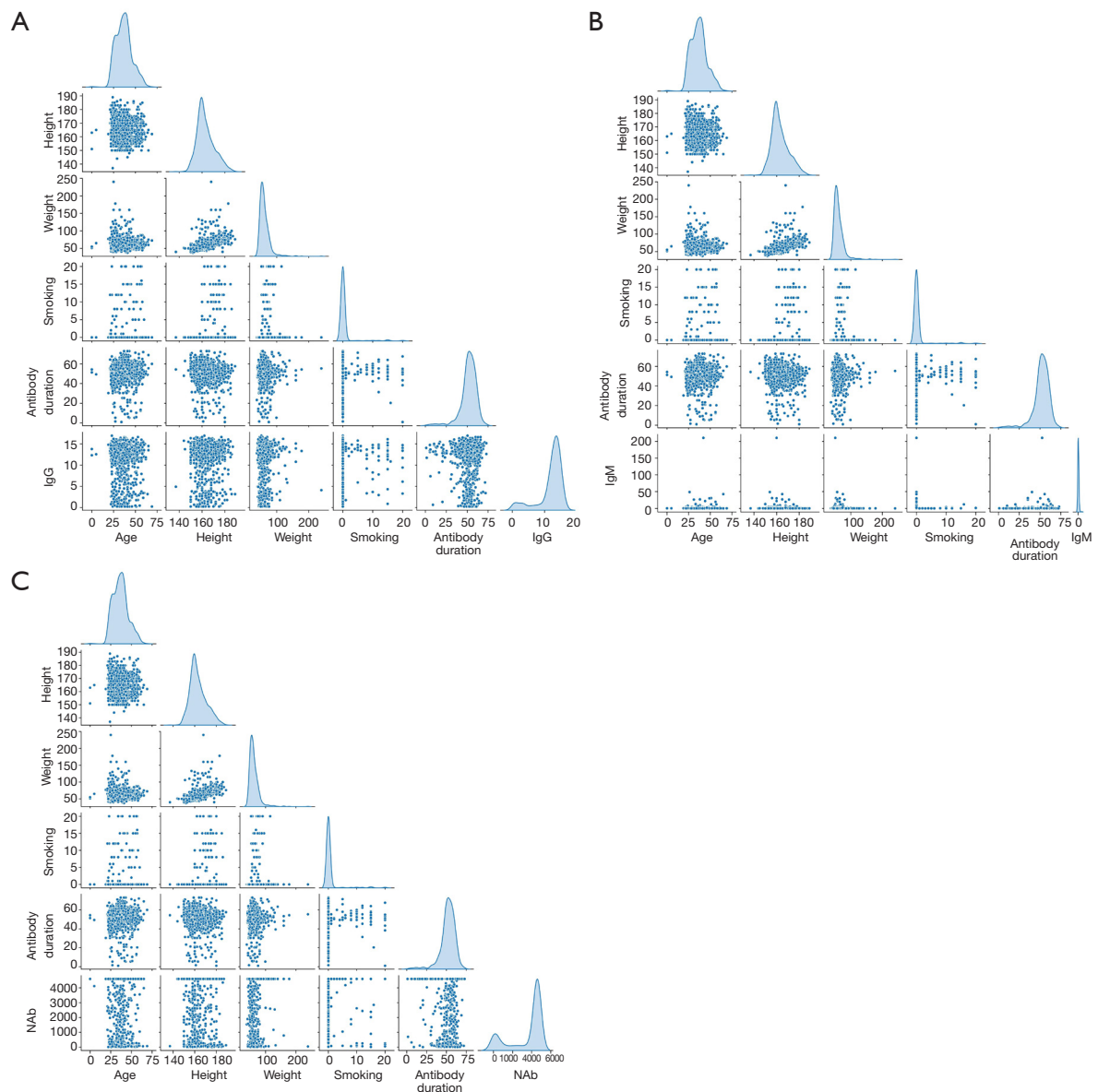
Figure 3 Distribution of personal properties of three types of antibodies. (A) The distribution of personal properties of IgG. The dots are in the scatter plot and at the edge are the distribution fitting plots, the same case in (B), which is for IgM, and (C), which is for NAb. IgG, immunoglobulin G; IgM, immunoglobulin M; NAb, neutralizing antibody.

Correlation of individual properties and antibody titer for COVID-19

The distribution of individual properties and the results of Spearman correlation analysis are presented in *Figures 3,4*. Notably, none of the individual properties exhibited strong correlations with each other (r<0.7), and regression coefficients (β), and 95% CIs for IgG, IgM, and NAb with individual baseline properties are provided in Figures S1-S3, then the

standardized regression coefficients (β) for these different antibodies are presented in Figures S4-S6. Consequently, all individual properties were included in the multilinear regression (MLR) model.

In the MLR model, we have included the following factors: age, height, weight, smoking status, and antibody duration. Model selection was based on the BIC and AIC scores. The model with the lowest score was deemed the optimized MLR model (*Table 3*). The analysis revealed that a history
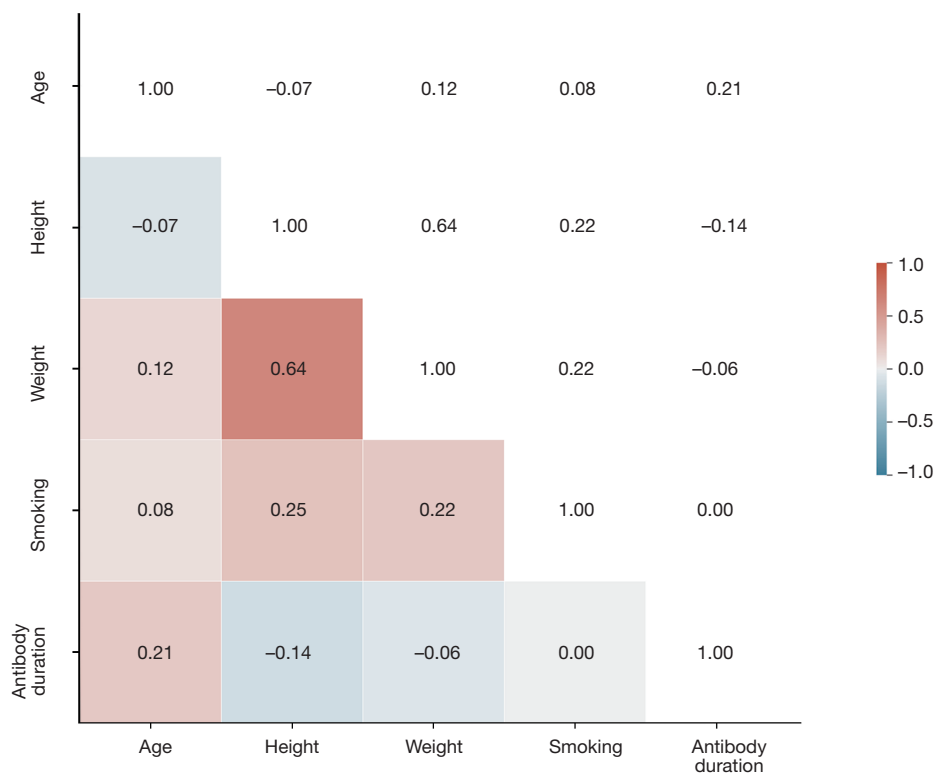
Figure 4 Spearman correlation matrix among factors other than three types of antibodies. Matrix is about correlation among factors other than three types of antibodies. The density of the color shows the strength of the correlation. Red is the positive correlation while blue is for negative correlation. Numbers on the matrix are the exact r value for each group.

of pulmonary disease and hospitalization had a significant impact on IgG titer. The optimized MLR model for IgG was determined to be $Y_{IgG} = f_0 + S(sex) + f(age) + f(height) + S(smk) + S(cd) + S(hosp) + S(his)$. However, none of the examined factors were found to be statistically significant for IgM titer, leading to the conclusion that $Y_{IgM} = f_0 + f(age)$ was the mixed MLR model. For NAb titers, we found that sex, smoking, family history of pulmonary disease, and hospitalization significantly influenced the titer. The optimized MLR model for NAb was determined to be $Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$.

For longitudinal analysis for the change of antibodies, optimized GBTM of IgG is an all-one-ordered four-group (1,1,1,1 group) model (*Figure 5*), average posterior probabilities for each group are 0.9735, 0.9665, 0.9894, and 08864, respectively. The results of the optimal model fitting showed that among the four categories, the first category group contained 52 individuals (4.3%) with low initial IgG

levels, followed by a gradual increase in IgG levels, which was defined as a "gradual growth type", and the other three groups contained 1,139 individuals (95.8%) with different initial IgG levels, followed by a gradual decrease in IgG levels, which was defined as a "gradual decline type". The study determined that the optimal IgM model is a full first-order group (1 group), and the average posterior probability of each group is 1.0000. The optimal model fitting results show that the initial IgM levels of each group are different, and then show a gradual decline, which is defined as a "gradual decline type". We determined that the optimal NAb model was also a full first-order group (1,1 group), and the average posterior probability of each group was 0.9790, and 0.9763, respectively. The results of fitting the optimal model showed that the initial NAb levels of each group varied, and then showed a gradual decline, which was defined as a "gradual decline type". Detailed calculation results of GBTM are provided in appendix available at https://cdn.amegroups.cn/static/public/jtd-23-1516-1.xlsx.

Simulation of the current COVID-19 epidemic by high-dimensional transmission dynamics model

As shown in *Figure 6*, in the absence of available data on the precise number of COVID-19 infections, we employed a high-dimensional transmission dynamics model that considers reinfection and the vaccination status of individuals to simulate the current epidemic in Xiamen City. Notably, after the initial peak observed in late December 2022 and January 2023, a second peak occurred in mid-May and early June in Xiamen City. Additionally, the model predicts a third peak in mid-October and early November. According to *Figure 7*, the analysis reveals variations in infection and incidence rates across different age groups. The age groups 30–39 and 40–49 years exhibit higher incidence and infection rates during both the second

Table 3 AIC and BIC results of optimized MLR model for three types of antibodies

Antibody	MLR model	AIC	BIC
IgG	$Y_{IgG} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$	6,580.65	8,501.69
	$Y_{IgG} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(hosp) + f(SI) + S(his)$	6,578.65	8,494.61
	$Y_{IgG} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(hosp) + S(his)$	6,576.66	8,487.75
	$Y_{IgG} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(PI) + S(hosp) + S(his)$	6,576.62	8,592.98
	$Y_{IgG} = f_0 + S(sex) + f(age) + S(loca) + f(height) + S(smk) + S(resd) + S(cd) + S(vac) + S(PI) + S(hosp) + S(his)$	6,574.67	8,586.66
	$Y_{IgG} = f_0 + S(sex) + f(age) + S(loca) + f(height) + S(smk) + S(resd) + S(cd) + S(vac) + S(hosp) + S(his)$	6,572.91	8,582.96
	$Y_{IgG} = f_0 + S(sex) + f(age) + S(loca) + f(height) + S(smk) + S(cd) + S(vac) + S(hosp) + S(his)$	6,571.19	8,579.77
	$Y_{IgG} = f_0 + S(sex) + f(age) + f(height) + S(smk) + S(cd) + S(vac) + S(hosp) + S(his)$	6,564.77	8,616.74
	$Y_{IgG} = f_0 + S(sex) + f(age) + f(height) + S(smk) + S(cd) + S(hosp) + S(his)$	6,561.42	8,611.84
	$Y_{IgG} = f_0 + S(sex) + f(age) + S(smk) + S(cd) + S(hosp) + S(his)$	6,560.03	8,613.38
	$Y_{IgG} = f_0 + S(sex) + S(smk) + S(cd) + S(hosp) + S(his)$	6,559.26	8,623.92
	$Y_{IgG} = f_0 + S(sex) + S(smk) + S(hosp) + S(his)$	6,559.28	8,645.66
	$Y_{IgG} = f_0 + S(smk) + S(hosp) + S(his)$	6,557.90	8,647.51

Table 3 (*continued*)

Table 3 (*continued*)

Antibody	MLR model	AIC	BIC
IgM	$Y_{IgM} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$	7,946.09	44,470.88
	$Y_{IgM} = f_0 + S(sex) + f(age) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$	7,935.07	44,471.87
	$Y_{IgM} = f_0 + S(sex) + f(age) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$	7,933.07	44,464.98
	$Y_{IgM} = f_0 + S(sex) + f(age) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$	7,925.57	44,540.34
	$Y_{IgM} = f_0 + S(sex) + f(age) + f(weight) + S(smk) + S(resd) + S(vac) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$	7,923.60	44,534.36
	$Y_{IgM} = f_0 + S(sex) + f(age) + f(weight) + S(smk) + S(resd) + S(vac) + S(PI) + S(symp) + f(SI) + S(his)$	7,921.63	44,528.56
	$Y_{IgM} = f_0 + S(sex) + f(age) + f(weight) + S(smk) + S(resd) + S(PI) + S(symp) + f(SI) + S(his)$	7,917.96	44,529.20
	$Y_{IgM} = f_0 + f(age) + f(weight) + S(smk) + S(resd) + S(PI) + S(symp) + f(SI) + S(his)$	7,916.00	44,523.82
	$Y_{IgM} = f_0 + f(age) + f(weight) + S(smk) + S(resd) + S(PI) + f(SI) + S(his)$	7,914.04	44,518.57
	$Y_{IgM} = f_0 + f(age) + f(weight) + S(resd) + S(PI) + f(SI) + S(his)$	7,912.11	44,514.47
	$Y_{IgM} = f_0 + f(age) + S(resd) + S(PI) + f(SI) + S(his)$	7,910.25	44,513.71
	$Y_{IgM} = f_0 + f(age) + S(resd) + f(SI) + S(his)$	7,908.85	44,533.38
	$Y_{IgM} = f_0 + f(age) + S(resd) + S(his)$	7,907.30	44,546.42
	$Y_{IgM} = f_0 + f(age) + S(resd)$	7,905.92	7,905.92
	$Y_{IgM} = f_0 + f(age)$	7,905.23	7,905.23

Table 3 (*continued*)

Table 3 (*continued*)

Antibody	MLR model	AIC	BIC
NAb	$Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(PI) + S(symp) + S(hosp) + f(SI) + S(his)$	21,148.52	3,437,296,211.85
	$Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(resd) + S(cd) + S(vac) + S(phsms) + S(symp) + S(hosp) + f(SI) + S(his)$	21,146.52	3,437,296,701.40
	$Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) + f(weight) + S(smk) + S(cd) + S(vac) + S(phsms) + S(symp) + S(hosp) + f(SI) + S(his)$	21,144.52	3,437,299,018.61
	$Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) + S(smk) + S(cd) + S(vac) + S(phsms) + S(symp) + S(hosp) + f(SI) + S(his)$	21,142.52	3,437,308,928.06
	$Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) + S(smk) + S(cd) + S(vac) + S(symp) + S(hosp) + f(SI) + S(his)$	21,138.11	3,453,456,677.99
	$Y_{NAb} = f_0 + S(sex) + f(age) + S(loca) + f(height) + S(smk) + S(cd) + S(vac) + S(symp) + S(hosp) + S(his)$	21,136.13	3,453,534,733.00
	$Y_{NAb} = f_0 + S(sex) + S(loca) + f(height) + S(smk) + S(cd) + S(vac) + S(symp) + S(hosp) + S(his)$	21,134.36	3,454,199,593.04
	$Y_{NAb} = f_0 + S(sex) + S(loca) + f(height) + S(smk) + S(vac) + S(symp) + S(hosp) + S(his)$	21,133.60	3,457,789,104.03
	$Y_{NAb} = f_0 + S(sex) + S(loca) + f(height) + S(smk) + S(vac) + S(hosp) + S(his)$	21,134.56	3,466,392,372.71
	$Y_{NAb} = f_0 + S(sex) + f(height) + S(smk) + S(vac) + S(hosp) + S(his)$	21,127.39	3,480,471,771.79
	$Y_{NAb} = f_0 + S(sex) + S(smk) + S(vac) + S(hosp) + S(his)$	21,128.19	3,488,662,304.51
	$Y_{NAb} = f_0 + S(sex) + S(smk) + S(hosp) + S(his)$	21,127.58	3,498,633,038.26

Loca, location/address; smk, whether smoke; resd, respiratory diseases history; cd, chronical diseases; vac, vaccination condition; phsms, pharmaceutical or non-pharmaceutical protection condition of individuals; PI, whether infected by SARS-CoV-2; symp, severity of symptoms; hosp, whether admitted in hospitals; SI, time interval between onset of symptoms and testing for antibody titer; his, family history of respiratory or pulmonary diseases. AIC, Akaike information criterion; BIC, Bayesian information criterion; MLR, multilinear regression; IgG, immunoglobulin G; IgM, immunoglobulin M; NAb, neutralizing antibody; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

and third peaks, followed by age groups 50–59, 18–29, and 70–79 years. Conversely, the eldest age group (80+) and the youngest [0–2] tend to have lower infection and incidence rates. However, it is essential to note that, overall, the incidence and infection rates for all age groups follow a similar scale as observed in the mid-aged group.

Discussion

The COVID-19 pandemic has sparked extensive research into prevention and symptom management (22). Notably, the quantity and longevity of antibodies produced during infection or vaccination are critical factors related to reinfection risk and symptom severity (23). The consensus
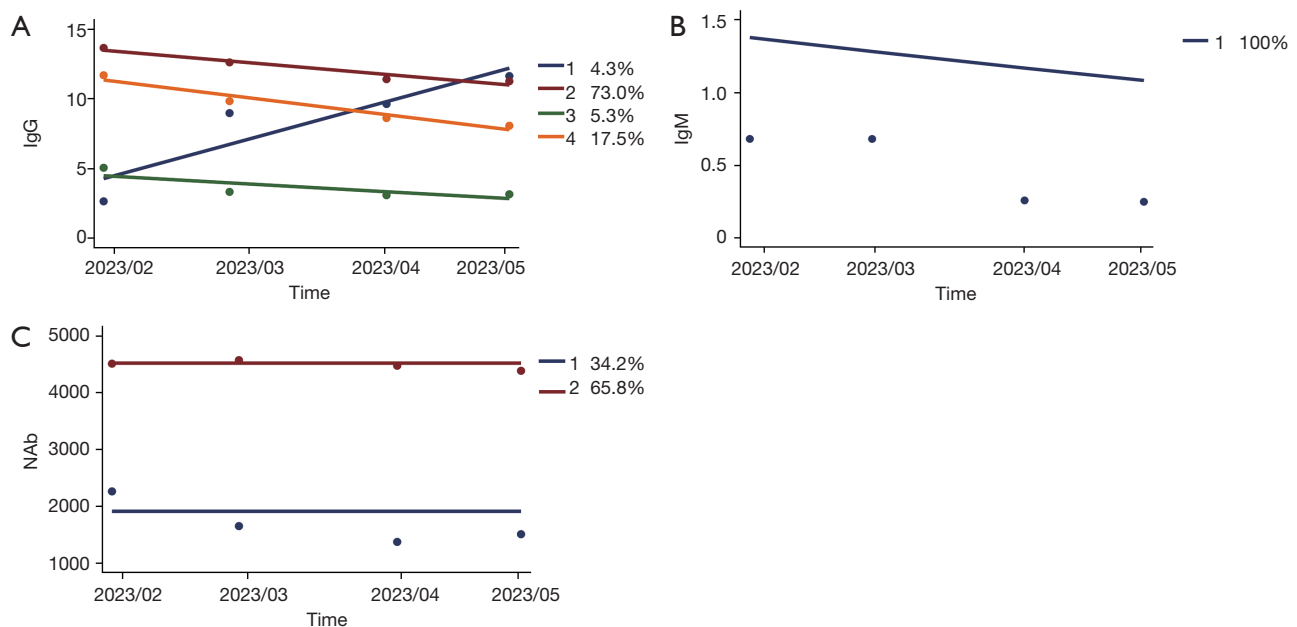
Figure 5 Result of GBTM on fitting the trend for three types of antibodies. (A) For IgG, 1, 2, 3, 4 represents four groups (1,1,1,1 group). 1 is the first category group. 4.3% is the proportion of the individuals that fit this group and it was an upward line, which was defined as a "gradual growth type", and the other three groups 2, 3, and 4 contained the rest of the individuals (95.8%) with different initial IgG levels, and showed a gradual decrease in IgG levels, which was defined as a "gradual decline type". Similarly, in (B), which is for IgM, there was only one line, showing that the optimal IgM model is a full first-order group (1 group), which is defined as a "gradual decline type". (C) For the NAb model, featuring two lines, which means it is also a full first-order group (1,1 group). IgG, immunoglobulin G; IgM, immunoglobulin M; NAb, neutralizing antibody; GBTM, group-based trajectory model.

holds that as antibody levels wane, individuals become more susceptible to reinfection, leading to discussions on the protective role of antibodies after natural infection or vaccination (24).

Crowd distribution for the samples

In this study, serum antibody titers of 1,191 public health workers in Xiamen City were tested over four consecutive monthly intervals. The study population exhibited no significant differences in antibody distribution due to a likely common immune response to both infection and vaccination (11). As the staff are all from the Xiamen Healthcare System, including doctors, CDCs workers, and primary healthcare organizations, it explains the insignificant difference in age groups and gender, but it also indicates that these indifferences could not be extrapolated to the general population.

However, for the differences in addresses, it is found that some staff live in cities near Xiamen and work in Xiamen

CDC, therefore, there would be some differences in the regional distribution of the staff. However, due to the limitations of the sample, we cannot assume that there is a difference in infection among the population in different districts of the whole of Xiamen City. What's more, before December 2022, it was recommended that people take one booster dose of the COVID-19 vaccine, yet after the new COVID-19 policy, there has been a new peak in people accepting vaccination (25). This questionnaire was sent out in February, therefore, there will be some differences in the vaccinated status between infected and non-infected groups.

Correlation of individual properties and antibody titer for COVID-19

Previously, there have been studies exploring the serological information for COVID-19 antibodies, which included immune process, seroconversion time, immune response as related to clinical presentation, and immune duration (3,5,6,10,26-29). Results of these studies have indicated
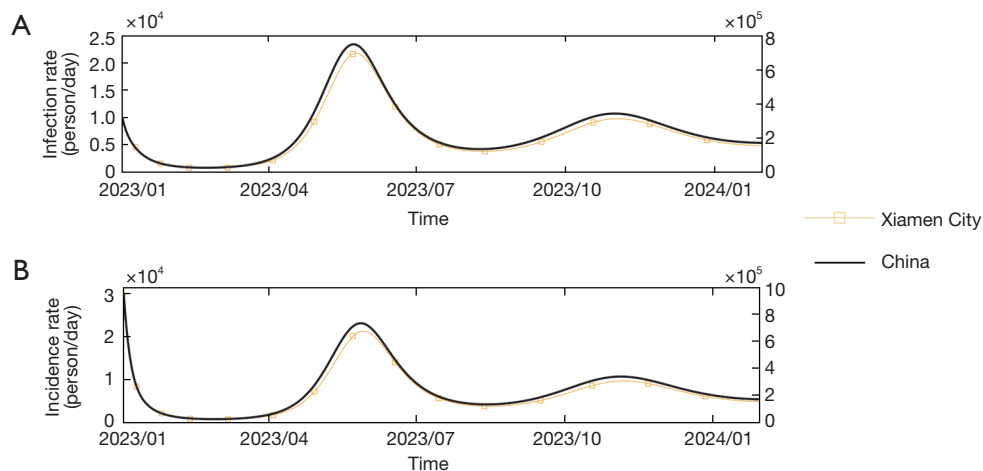
Figure 6 Simulated infection rate and incidence rate of COVID-19 in Xiamen City and China. (A) The figure for the daily infection rate, X-axis shows the prediction time. The left Y-axis is for Xiamen City and the right one is for China, which is the same in (B) for the daily incidence rate. COVID-19, coronavirus disease 2019.
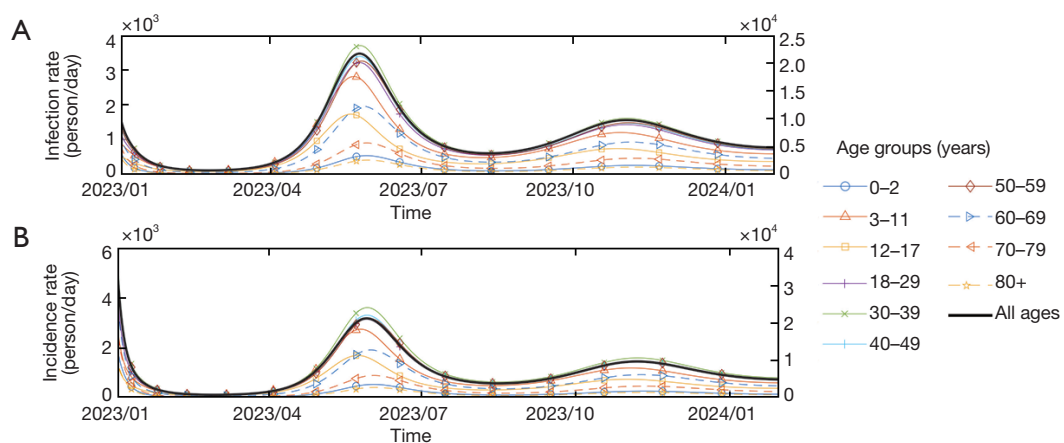
Figure 7 Simulated infection rate and incidence rate of COVID-19 for different age groups in Xiamen City. (A) The figure for the daily infection rate, and the X-axis shows the prediction time. The left Y-axis is for various age groups and the right one is for all ages, which is the same in (B) of daily incidence rate. COVID-19, coronavirus disease 2019.

that immune responses to SARS-CoV-2 are consistent with general viral infection patterns. However, some researchers suggest that immunity generated from natural infection appears to be short, which gives a hint that reinfection is inevitable (6,30-32). Should we find out individual properties that could make a difference in the antibody titer, it would be possible for us to select those who are at a high risk of short-term reinfection.

A MLR analysis was conducted to identify factors influencing IgG, IgM, and NAb titers. For IgM, only age demonstrated a significant impact, likely due to its role as

the primary response and potential age-related variations in immune response (5); IgM plays a prominent role during the early age and only exists for 20–30 days. Since T-cell-derived antibody production decreases and B-lymphocyte generation decreases with age, antibody response against infectious agents and after vaccination may not be sufficient (17). In contrast, IgG, whose titers remained elevated and relatively stable for a longer period after induced, were significantly influenced by age, weight, height, history of pulmonary disease, and hospitalization. History of pulmonary diseases could be illustrated as soundness of

the immune system, which indicates that those who have suffered from pulmonary diseases may have pulmonary immune system damage and result in lower antibody titers (18). Hospitalization could be explained as the severity of symptoms, which can be indicative of immune system health and symptom severity (19). The effect of smoking on antibody response has shown variable results with different viruses. For example, a very low antibody response was measured in smokers after the hepatitis B vaccine (20). NAb titers were impacted by sex, smoking, family history of pulmonary disease, and hospitalization. Smoking's effect on NAb titers can be explained by NAb's longer-lasting nature, indicating its significance in protection (8).

Ever since the pandemic, there have been a variety of studies that revealed the duration of different antibodies through typical textbook knowledge as well as experimental studies, cohort studies, etc. (5,21,29,33). Here in this study, we have collected and tested for serum antibodies of 1,191 healthcare workers four times and applied a GBTM model to explore potential development trajectories of antibodies. It was found in this study that only IgG presented a different trajectory, while IgM and NAb are the unchanged type. As IgM indicates the instant infections, remained unchanged since the timing for serological tests were all conducted in the middle of the month, when the infection was over. For NAb, it may be because they could last for more than 3 months, and due Labor Day holiday from May 1st to 5th, people started travelling and increased the possibility of contact, then there was a peak in the infection rate, which may be the reason why NAb titer was unchanged. However, it was found that there were 52 individuals had a low IgG titer at first then kept growing during the follow-up period, while other 1,139 individuals with different initial IgG titer kept decreasing. Then we made a further investigation on those 52 individuals and found out that there were no differences in their infection status at the first serum test or in their gender, age, and other properties. Therefore, we need a further exploration of the reason why some would show an increasing trend while others declined (34).

Simulation of the current COVID-19 epidemic by high-dimensional transmission dynamics model

Here we are under the circumstance that obtaining the current epidemic is unable to be acquired. We refined the previous SEIAR (Susceptible-Exposed-Infection-Asymptomate-Removed/Recovered) transmission dynamics model (35), we have introduced some real-world scenarios

such as vaccination, pre-symptomatic infections, reinfection to the simple model and constructed VEAFIR$_p$RV model. Then we simulated the future incidence rate and infection rate of COVID-19. We have simulated and predicted that there was a peak in the middle of May and early July here in Xiamen City, which may result from an increased contact rate among the population due to the Labor Day holiday and Dragonboat Festival. And it was consistent with the time when IgG titer decreased and NAb titer increased in the tested participants. Hence, we believe that the authorities should remind people to take precautions during holidays when there is a large movement of people, and the government should be prepared for the possible shortage of medical resources due to COVID-19 during holidays.

Then we predicted that starting from late September, there would be another peak in the mid-October to early November. This could be explained by the increase in contact due to school starting, National Holidays, and cold weather, leading to low ventilation rates in rooms. Therefore, there may be a further decrease in IgG and another increase in NAb. Another prediction and simulation of incidence rate and infection rate was for different age groups. The quality of the antibody response is influenced by age, as is the case for T-cell responses (17). Given the risk of severe COVID-19 in the elderly is higher, age-dependent immunological mechanisms are particularly important to elucidate, as these could be targeted to improve responses to natural infection and vaccination (29). During simulation, we also found that the highest infection and incidence rate lie in the age group of 30–39 years, where the old and the babies are the lowest, we believe this result was driven by the total population of Xiamen City and the number of people by age groups. The independent population (age 15–60 years) has more chance of contact with each other than the dependent population (age <15 and >60 years).

Limitation

The limitation of the study is that the participants did not reflect the general population, as the study was conducted in a small group of healthy volunteers and we applied convenience sampling. Therefore, we could not adequately examine the differences in antibody levels in the presence of various diseases that are more likely to be present in the population with the demographic data we examine. Another limitation of this study lies in the high-dimensional transmission dynamics model. In this model, the contact among the population is considered to be the major factor

affecting the epidemic peak, mutation of the virus as well as the possible immune escape of variants are not taken into consideration, which indicates fact that the model is supposed to be refined to make a better prediction on the future epidemic trend.

Conclusions

While COVID-19 policies have evolved, the health challenges posed by the disease remain significant. New mutants continue to emerge, suggesting that reinfection is still a pressing problem. This study reveals key information about COVID-19 antibody dynamics in Xiamen healthcare workers. It highlights age-related impacts on IgM and diverse influences on IgG, and NAb titers, emphasizing the complexity of individual characteristics in antibody responses. However, the distinct trajectory in IgG levels among subsets warrants further investigation for potential factors driving these contrasting trends.

To build a more robust defense against reinfection and optimize self-protection strategies, we must delve deeper into understanding antibody dynamics and tailor more effective public health strategies as we navigate the ongoing challenges posed by COVID-19 epidemics.

Acknowledgments

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Jing Cheng, Tao Xu, Zifeng Yang, Wenda Guan) for the series "Current Status of Diagnosis and Forecast of COVID-19" published in *Journal of Thoracic Disease*. The article has undergone external peer review.

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-1516/rc

Data Sharing Statement: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-1516/dss

Peer Review File: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-1516/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jtd.amegroups.com/article/view/10.21037/jtd-23-1516/coif). The series "Current Status of Diagnosis and Forecast of COVID-19" was commissioned by the editorial office without any funding or sponsorship. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of the Xiamen Center for Disease Control [XJK/LLSC(2022)004] and participants signed the agreement consent before filling out the questionnaire.

References

1. World Health Organization. Strategy and planning. Available online: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/strategies-and-plans
2. Cohen JI, Burbelo PD. Reinfection With SARS-CoV-2: Implications for Vaccines. Clin Infect Dis 2021;73:e4223-8.
3. Castro Dopico X, Ols S, Loré K, et al. Immunity to SARS-CoV-2 induced by infection or vaccination. J Intern Med 2022;291:32-50.
4. Cruz AT, Zeichner SL. Duration of Effective Antibody Levels After COVID-19. Pediatrics 2021;148:e2021052589.
5. Fearon DT, Locksley RM. The instructive role of innate immunity in the acquired immune response. Science 1996;272:50-3.
6. Sneller MC, Liang CJ, Marques AR, et al. A Longitudinal

Study of COVID-19 Sequelae and Immunity: Baseline Findings. Ann Intern Med 2022;175:969-79.

7. Jiang JC, Zhang Y. Serological antibody testing in the COVID-19 pandemic: their molecular basis and applications. Biochem Soc Trans 2020;48:2851-63.

8. Evans JP, Zeng C, Carlin C, et al. Neutralizing antibody responses elicited by SARS-CoV-2 mRNA vaccination wane over time and are boosted by breakthrough infection. Sci Transl Med 2022;14:eabn8057.

9. De Giorgi V, West KA, Henning AN, et al. Naturally Acquired SARS-CoV-2 Immunity Persists for Up to 11 Months Following Infection. J Infect Dis 2021;224:1294-304.

10. Padoan A, Dall'Olmo L, Rocca FD, et al. Antibody response to first and second dose of BNT162b2 in a cohort of characterized healthcare workers. Clin Chim Acta 2021;519:60-3.

11. Nordström P, Ballin M, Nordström A. Risk of SARS-CoV-2 reinfection and COVID-19 hospitalisation in individuals with natural and hybrid immunity: a retrospective, total population cohort study in Sweden. Lancet Infect Dis 2022;22:781-90.

12. Hachmann NP, Miller J, Collier AY, et al. Neutralization escape by SARS-CoV-2 Omicron subvariants BA. 2.12. 1, BA. 4, and BA. 5. N Engl J Med 2022;387:86-8.

13. Brand SPC, Ojal J, Aziza R, et al. COVID-19 transmission dynamics underlying epidemic waves in Kenya. Science 2021;374:989-94.

14. Cai J, Deng X, Yang J, et al. Modeling transmission of SARS-CoV-2 Omicron in China. Nat Med 2022;28:1468-75.

15. Donnat C, Bunbury F, Kreindler J, et al. Predicting COVID-19 Transmission to Inform the Management of Mass Events: Model-Based Approach. JMIR Public Health Surveill 2021;7:e30648.

16. Memoli V, Ekanmian G, Lunghi C, et al. What methods are used to study the association between medication adherence trajectories, estimated with the group-based trajectory modeling (GBTM) method, and health-related outcomes?-a protocol for a systematic review. Syst Rev 2022;11:102.

17. Grifoni A, Weiskopf D, Ramirez SI, et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. Cell 2020;181:1489-1501.e15.

18. Swartz MD, DeSantis SM, Yaseen A, et al. Antibody Duration After Infection From SARS-CoV-2 in the Texas Coronavirus Antibody Response Survey. J Infect Dis 2023;227:193-201.

19. Smits VAJ, Hernández-Carralero E, Paz-Cabrera MC, et al. The Nucleocapsid protein triggers the main humoral immune response in COVID-19 patients. Biochem Biophys Res Commun 2021;543:45-9.

20. Huang YP, Gauthey L, Michel M, et al. The relationship between influenza vaccine-induced specific antibody responses and vaccine-induced nonspecific autoantibody responses in healthy older women. J Gerontol 1992;47:M50-5.

21. Piccoli L, Ferrari P, Piumatti G, et al. Risk assessment and seroprevalence of SARS-CoV-2 infection in healthcare workers of COVID-19 and non-COVID-19 hospitals in Southern Switzerland. Lancet Reg Health Eur 2021;1:100013.

22. Iwasaki A, Yang Y. The potential danger of suboptimal antibody responses in COVID-19. Nat Rev Immunol 2020;20:339-41.

23. Newell KL, Clemmer DC, Cox JB, et al. Switched and unswitched memory B cells detected during SARS-CoV-2 convalescence correlate with limited symptom duration. PLoS One 2021;16:e0244855.

24. Zhang Y, Zeng G, Pan H, et al. Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine in healthy adults aged 18-59 years: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial. Lancet Infect Dis 2021;21:181-92.

25. Nair S, Chen X. Biology of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the humoral immunoresponse: a systematic review of evidence to support global policy-level actions and research. Glob Health J 2022;6:38-43.

26. Baumgarth N, Nikolich-Žugich J, Lee FE, et al. Antibody Responses to SARS-CoV-2: Let's Stick to Known Knowns. J Immunol 2020;205:2342-50.

27. Dan JM, Mateus J, Kato Y, et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. Science 2021;371:eabf4063.

28. Lee WS, Wheatley AK, Kent SJ, et al. Antibody-dependent enhancement and SARS-CoV-2 vaccines and therapies. Nat Microbiol 2020;5:1185-91.

29. Xie J, Ding C, Li J, et al. Characteristics of patients with coronavirus disease (COVID-19) confirmed using an IgM-IgG antibody test. J Med Virol 2020;92:2004-10.

30. Kusunoki H, Ohkusa M, Iida R, et al. Longitudinal Changes in IgG-Type SARS-CoV-2 Antibody Titers after COVID-19 Vaccination and a Prominent Increase in Antibody Titers When Infected after Vaccination. Vaccines

(Basel) 2023;11:860.

31. Chen J, Liu X, Zhang X, et al. Decline in neutralising antibody responses, but sustained T-cell immunity, in COVID-19 patients at 7 months post-infection. Clin Transl Immunology 2021;10:e1319.

32. Iyer AS, Jones FK, Nodoushani A, et al. Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. Sci Immunol 2020;5:eabe0367.

33. Uysal EB, Gümüş S, Bektöre B, et al. Evaluation of antibody response after COVID-19 vaccination of healthcare workers. J Med Virol 2022;94:1060-6.

34. Altawalah H. Antibody Responses to Natural SARS-CoV-2 Infection or after COVID-19 Vaccination. Vaccines (Basel) 2021;9:910.

35. Niu Y, Rui J, Wang Q, et al. Containing the Transmission of COVID-19: A Modeling Study in 160 Countries. Front Med (Lausanne) 2021;8:701836.

Appendix 1

Methods

Data collection

Data used in this study includes COVID-19 histories of individuals and a follow-up serum antibody test. As it is more available to conduct this research among healthcare workers, convenience sampling is a practical way to collect data. However, as participants are chosen based on ease of access, it may not represent the broader population accurately.

First, we initiated data collection by distributing questionnaires to public health workers in Xiamen healthcare system, including hospitals, CDCs, and primary healthcare organizations, and questionnaires were distributed to samples who were selected via convenient sampling. As we targeted individuals who had been infected by SARS-CoV-2 or had received vaccinations after December 1st, 2022. After collecting and reviewing the completed questionnaires, a total of 1,344 questionnaires were collected. Based on prior research indicating that NAbs for COVID-19 might persist for a maximum of 73 days (34), questionnaires that were incomplete or lacking essential infection-related information were excluded from the study. Finally, 1,191 of them being eligible to be included as the study's sample population.

Subsequently, we conducted serum antibody testing on the selected participants in four monthly intervals, specifically in February, March, April, and May. This longitudinal testing aimed to capture the variations in antibody titers over time.

MLR models

MLR models were established to analyze the relationship between individual baseline properties and COVID-19 antibody titers. Regression coefficients (β), 95% CIs, and standardized regression coefficients (β) are calculated and the calculation results are presented in *Figures S1-S6*.

GBTMs

The study constructs a GBTM, which can depict the characteristic dynamic changes of time-varying variables as the number of follow-up visits increases. It simultaneously divides the population into several latent class groups and establishes a latent growth class model within each category to describe the individual changes over time within the group. This model can not only reveal the relationships between different latent trajectories but also depict the fluctuations within the trajectory, thus providing a more realistic grouping of indicators and conducting predictive research (16).

The study uses the "traj" package in Stata 17.0 software for data analysis, first analyzing the dynamic changes in serum IgG, IgM, and NAb from baseline, the first to the third follow-up visits, and employing GBTM to identify latent clusters with similar trajectories. The study hypothesizes that the population serum may be divided into up to five main categories: gradual growth type, gradual decline type, unchanged type, growth and then decline type, decline and then growth type. The model is set with the highest order of 3, meaning the potential groupings of the population are between 1 and 5 groups, with orders 0–3, where the order reflects the speed of the trend changes. The model's effectiveness is evaluated using the BIC and the AIC, with the smallest absolute value of the indicators and closer to 0 indicating a better model fit. Additionally, a higher average posterior probability of group membership (AvePP) indicates a better model fit. To ensure the effectiveness of the grouping, each group composition in the model should account for at least 2–5% and must be consistent with medical knowledge.

After determining the GBTM groupings, χ^2 analysis is used to compare differences in serological change patterns by gender, age, and region, and analysis of variance is used to compare differences in serological change patterns by gender, age, and region. Multivariate linear regression and multivariate logistic regression are used to compare demographic differences between different serological change patterns. A P value of <0.05 is considered statistically significant. See the GBTM results of IgG, IgM and NAb in appendix available at https://cdn.amegroups.cn/static/public/jtd-23-1516-1.xlsx.

High-dimensional transmission dynamics model

VEAFIR$_p$RV model was built under these assumptions:

In this model, we grouped the total population N into susceptible, V; exposed, E; asymptomatic infection, A; pre-symptomatic infection, F; symptomatic infection, I; removed/recovered that will not be reinfected, R$_p$; recovered and will possible be reinfected, R.

(I) First, for the total population, we have included various regions (i), age groups (j) and their vaccination status (k) in the compartment N when making the calculation.

(II) Transmission rate, which is in the dimension of [person·time^{-1}], contributes to the reduction of susceptible population (V) and the increase of exposed population (E). Usually, in an ODE model, it simplifies individual properties to the population averaged quantities, and

depicts the age-heterogeneity of transmission by the 'contact frequency matrix'. Therefore, we assumed that should age-specific contact frequency matrices for various regions are available, then for every individual in different age groups, the expected number of symptomatic infectious individuals he/she has contacted during a time interval could be calculated. Then the vaccine efficacy, which considered to be able to reduce transmissibility, was included in to calculation.

(III) In an ODE model, any infected individual is first categorized as exposed (E), then at time t, according to the natural history of COVID-19, there would be two results for exposed population E, they either become pre-symptomatic infection F or asymptomatic infections A. Assume that a proportion p of E is converted to A, and the proportion of E to I is (1 – p). It is generally believed that after a person is exposed to pathogens, the time interval of he/she gets invaded by the pathogen and able to emit it is called latent period. The rate of transformation from E to A is proportional to the amount of E with a scale factor of $p\omega E$ and ω is the latent period coefficient. And in the case of symptomatic infections, since there would be a time lag between the time of virus excretion and symptoms onset, we set ω' and ω'' as average incubation period from exposed (E) to pre-symptomatic infection (F) and average incubation period from pre-symptomatic infection (F) to infections (I), respectively.

(IV) After infection, individuals would be removed/recovered, however, some of the removed/recovered population could experience reinfection while others won't. Therefore, we set two endings for COVID-19 infections: one is the removed/recovered group that would not be infected again (R_p), and the other is removed/recovered group that would experience reinfection (R). Here we set the proportion of those become susceptible again after recovery as δ. At time t, the number of transfers to R and R_p is γI if the time interval between onset and diagnosis from a symptomatic infection I is γ; the number of transfers to R from A who is identified as asymptomatic infection is $\gamma' I$.

(V) As reinfection is often correlated with losing immunity, therefore, for R_p become susceptible, V depends on the immunity duration of the individual, here we introduce τ to be the average duration coefficient for immunity duration at time t.

Sensitivity analysis

This High-dimensional transmission dynamics model is an extension of basic SEIAR model, or we could consider it as an SEIAR model with multiple groups. Since the SEIAR models with multi-group are widely used by many studies (36,37), the sensitivity of other model parameters could be found in those references. In our model, the vector of VE, is multiplied on the group-wise contact matrix; which makes the sensitivity analysis analogous to those for the contact matrix.

Additional results

Dep. Variable:	IgG	No. Observations:	1191
Model:	GLM	Df Residuals:	1165
Model Family:	Gaussian	Df Model:	25
Link Function:	identity	Scale:	14.380
Method:	IRLS	Log-Likelihood:	-3264.3
Date:	Thu, 31 Aug 2023	Deviance:	16753.
Time:	08:59:39	Pearson chi2:	1.68e+04
No. Iterations:	3	Pseudo R-squ. (CS):	0.03206
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	15.7720	4.890	3.225	0.001	6.187	25.357
sex[T.Female]	-0.4605	0.377	-1.221	0.222	-1.200	0.279
resd[T.Yes]	-0.1378	0.270	-0.511	0.610	-0.667	0.391
cd[T.Yes]	0.4007	0.410	0.978	0.328	-0.403	1.204
loca[T.XiangAn district]	-1.7410	2.796	-0.623	0.533	-7.220	3.738
loca[T.TongAn district]	-1.6696	2.793	-0.598	0.550	-7.145	3.806
loca[T.HaiCang district]	-2.2075	2.802	-0.788	0.431	-7.698	3.283
loca[T.JiMei district]	-1.4597	2.783	-0.525	0.600	-6.914	3.995
loca[T.HuLi district]	-2.3232	2.787	-0.834	0.404	-7.785	3.139
loca[T.SiMing district]	-1.9249	2.772	-0.694	0.487	-7.358	3.509
vac[T.Basic immnue]	0.8088	1.071	0.755	0.450	-1.291	2.908
vac[T.Booster immnue]	0.4963	0.921	0.539	0.590	-1.309	2.302
phsms[T.very low]	1.0027	0.889	1.128	0.259	-0.740	2.745
phsms[T.low]	-0.2880	1.055	-0.273	0.785	-2.356	1.780
phsms[T.mid]	0.3893	0.878	0.443	0.657	-1.331	2.110
phsms[T.high]	0.8134	0.864	0.942	0.346	-0.880	2.506
phsms[T.very high]	-0.1828	0.985	-0.186	0.853	-2.114	1.748
PI[T.Yes]	-0.1824	0.671	-0.272	0.786	-1.497	1.133
symp[T.Yes]	0.0122	0.608	0.020	0.984	-1.180	1.204
hosp[T.Yes]	-1.4475	0.658	-2.201	0.028	-2.736	-0.159
fam[T.Yes]	0.8821	0.270	3.270	0.001	0.353	1.411
age	-0.0123	0.014	-0.898	0.369	-0.039	0.015
height	-0.0194	0.023	-0.840	0.401	-0.065	0.026
weight	0.0016	0.009	0.186	0.852	-0.015	0.018
smk	-0.0558	0.043	-1.309	0.190	-0.139	0.028
SI	-0.0016	0.013	-0.124	0.901	-0.026	0.023

Figure S1 Regression coefficients (β), 95% CIs for IgG with individual baseline properties. CI, confidence interval; IgG, immunoglobulin G.

```
Intercept                        12.23
sex[T.Female]                    -4.63
resd[T.Yes]                      -1.94
cd[T.Yes]                         3.71
loca[T.XiangAn district]         -2.36
loca[T.TongAn district]          -2.27
loca[T.HaiCang district]         -2.99
loca[T.JiMei district]           -1.99
loca[T.HuLi district]            -3.16
loca[T.SiMing district]          -2.63
vac[T.Basic immnue]               2.86
vac[T.Booster immnue]             2.04
phsms[T.very low]                 4.28
phsms[T.low]                     -1.03
phsms[T.mid]                      1.68
phsms[T.high]                     3.57
phsms[T.very high]               -0.70
PI[T.Yes]                        -1.03
symp[T.Yes]                       0.08
hosp[T.Yes]                      -8.35
fam[T.Yes]                       12.40
age                              -3.40
height                           -3.19
weight                            0.71
smk                              -4.96
SI                               -0.47
dtype: float64
```

Figure S2 Standardized regression coefficients (β) for IgG with individual baseline properties. IgG, immunoglobulin G.

Dep. Variable:	IgM	No. Observations:	1191
Model:	GLM	Df Residuals:	1165
Model Family:	Gaussian	Df Model:	25
Link Function:	identity	Scale:	45.255
Method:	IRLS	Log-Likelihood:	-3947.0
Date:	Thu, 31 Aug 2023	Deviance:	52722.
Time:	09:04:23	Pearson chi2:	5.27e+04
No. Iterations:	3	Pseudo R-squ. (CS):	0.008053
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-0.2712	8.675	-0.031	0.975	-17.275	16.732
sex[T.Female]	0.1216	0.669	0.182	0.856	-1.190	1.433
resd[T.Yes]	0.5441	0.479	1.137	0.256	-0.394	1.483
cd[T.Yes]	-0.0659	0.727	-0.091	0.928	-1.491	1.359
loca[T.XiangAn district]	-0.7255	4.959	-0.146	0.884	-10.445	8.994
loca[T.TongAn district]	-0.5136	4.956	-0.104	0.917	-10.226	9.199
loca[T.HaiCang district]	-0.3873	4.970	-0.078	0.938	-10.128	9.354
loca[T.JiMei district]	-0.2392	4.937	-0.048	0.961	-9.915	9.437
loca[T.HuLi district]	-0.7693	4.944	-0.156	0.876	-10.459	8.920
loca[T.SiMing district]	-0.2762	4.918	-0.056	0.955	-9.915	9.363
vac[T.Basic immnue]	0.3385	1.900	0.178	0.859	-3.386	4.063
vac[T.Booster immnue]	0.6315	1.634	0.387	0.699	-2.571	3.834
phsms[T.very low]	-0.3141	1.577	-0.199	0.842	-3.405	2.777
phsms[T.low]	1.2639	1.872	0.675	0.500	-2.405	4.932
phsms[T.mid]	-0.1162	1.557	-0.075	0.941	-3.169	2.936
phsms[T.high]	0.2525	1.532	0.165	0.869	-2.751	3.256
phsms[T.very high]	-0.2891	1.748	-0.165	0.869	-3.715	3.136
PI[T.Yes]	-0.6548	1.190	-0.550	0.582	-2.987	1.678
symp[T.Yes]	0.2038	1.079	0.189	0.850	-1.910	2.318
hosp[T.Yes]	-0.1212	1.166	-0.104	0.917	-2.407	2.165
fam[T.Yes]	0.3881	0.478	0.811	0.417	-0.550	1.326
age	0.0322	0.024	1.322	0.186	-0.016	0.080
height	0.0027	0.041	0.067	0.947	-0.078	0.083
weight	-0.0029	0.015	-0.193	0.847	-0.033	0.027
smk	-0.0143	0.076	-0.190	0.850	-0.163	0.134
SI	-0.0137	0.023	-0.607	0.544	-0.058	0.031

Figure S3 Regression coefficients (β), 95% CIs for IgM with individual baseline properties. CI, confidence interval; IgM, immunoglobulin M.

```
Intercept                     -0.21
sex[T.Female]                  1.22
resd[T.Yes]                    7.65
cd[T.Yes]                     -0.61
loca[T.XiangAn district]      -0.98
loca[T.TongAn district]       -0.70
loca[T.HaiCang district]      -0.52
loca[T.JiMei district]        -0.33
loca[T.HuLi district]         -1.05
loca[T.SiMing district]       -0.38
vac[T.Basic immnue]            1.20
vac[T.Booster immnue]          2.60
phsms[T.very low]             -1.34
phsms[T.low]                   4.54
phsms[T.mid]                  -0.50
phsms[T.high]                  1.11
phsms[T.very high]            -1.11
PI[T.Yes]                     -3.70
symp[T.Yes]                    1.27
hosp[T.Yes]                   -0.70
fam[T.Yes]                     5.46
age                            8.90
height                         0.45
weight                        -1.30
smk                           -1.27
SI                            -4.08
dtype: float64
```

Figure S4 Standardized regression coefficients (β) for IgM with individual baseline properties. IgM, immunoglobulin M.

Dep. Variable:	nAb	No. Observations:	1191
Model:	GLM	Df Residuals:	1165
Model Family:	Gaussian	Df Model:	25
Link Function:	identity	Scale:	2.9505e+06
Method:	IRLS	Log-Likelihood:	-10548.
Date:	Thu, 31 Aug 2023	Deviance:	3.4373e+09
Time:	09:06:53	Pearson chi2:	3.44e+09
No. Iterations:	3	Pseudo R-squ. (CS):	0.04587
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	6737.8837	2215.164	3.042	0.002	2396.243	1.11e+04
sex[T.Female]	-445.4320	170.808	-2.608	0.009	-780.210	-110.654
resd[T.Yes]	3.3422	122.250	0.027	0.978	-236.264	242.949
cd[T.Yes]	160.6590	185.645	0.865	0.387	-203.198	524.516
loca[T.XiangAn district]	-1883.5353	1266.276	-1.487	0.137	-4365.391	598.321
loca[T.TongAn district]	-1678.3478	1265.352	-1.326	0.185	-4158.392	801.697
loca[T.HaiCang district]	-2064.9435	1269.009	-1.627	0.104	-4552.156	422.268
loca[T.JiMei district]	-1823.9418	1260.520	-1.447	0.148	-4294.516	646.632
loca[T.HuLi district]	-1941.9523	1262.277	-1.538	0.124	-4415.971	532.066
loca[T.SiMing district]	-1826.6031	1255.714	-1.455	0.146	-4287.758	634.551
vac[T.Basic immnue]	1001.7013	485.227	2.064	0.039	50.674	1952.729
vac[T.Booster immnue]	768.6595	417.213	1.842	0.065	-49.062	1586.381
phsms[T.very low]	329.6995	402.675	0.819	0.413	-459.530	1118.929
phsms[T.low]	-301.3478	477.924	-0.631	0.528	-1238.063	635.367
phsms[T.mid]	79.2618	397.682	0.199	0.842	-700.180	858.704
phsms[T.high]	198.8071	391.255	0.508	0.611	-568.039	965.653
phsms[T.very high]	96.3143	446.271	0.216	0.829	-778.360	970.989
PI[T.Yes]	-3.9426	303.887	-0.013	0.990	-599.550	591.665
symp[T.Yes]	242.3555	275.415	0.880	0.379	-297.449	782.160
hosp[T.Yes]	-657.7967	297.845	-2.209	0.027	-1241.562	-74.032
fam[T.Yes]	559.6020	122.176	4.580	0.000	320.141	799.063
age	-3.0520	6.223	-0.490	0.624	-15.250	9.146
height	-16.3906	10.478	-1.564	0.118	-36.928	4.147
weight	-0.2215	3.892	-0.057	0.955	-7.849	7.406
smk	-54.6772	19.321	-2.830	0.005	-92.545	-16.809
SI	1.3280	5.759	0.231	0.818	-9.959	12.615

Figure S5 Regression coefficients (β), 95% CIs for NAb with individual baseline properties. CI, confidence interval; NAb, neutralizing antibody.

```
Intercept                     5,224.73
sex[T.Female]                -4,479.39
resd[T.Yes]                      46.96
cd[T.Yes]                     1,486.51
loca[T.XiangAn district]     -2,555.00
loca[T.TongAn district]      -2,278.33
loca[T.HaiCang district]     -2,795.05
loca[T.JiMei district]       -2,485.46
loca[T.HuLi district]        -2,642.59
loca[T.SiMing district]      -2,498.62
vac[T.Basic immnue]           3,546.00
vac[T.Booster immnue]         3,164.63
phsms[T.very low]             1,406.40
phsms[T.low]                 -1,083.07
phsms[T.mid]                    342.35
phsms[T.high]                   872.81
phsms[T.very high]              370.71
PI[T.Yes]                       -22.29
symp[T.Yes]                   1,511.51
hosp[T.Yes]                  -3,793.57
fam[T.Yes]                    7,867.55
age                            -842.37
height                       -2,686.88
weight                          -97.74
smk                          -4,861.00
SI                              396.08
dtype: float64
```

Figure S6 Standardized regression coefficients (β) for NAb with individual baseline properties. NAb, neutralizing antibody.

References

36. Zhao Z, Chen Q, Zhao B, et al. Transmission pattern of shigellosis in Wuhan City, China: a modelling study. Epidemiology & Infection 2021;149:e249.

37. Zhao ZY, Zhu YZ, Xu JW, et al. A five-compartment model of age-specific transmissibility of SARS-CoV-2. Infect Dis Poverty 2020;9:117.