Paradoxical dynamics of SARS-CoV-2 by herd immunity and antibody-dependent enhancement

Yasuhiko Kamikubo[1,4] *, Atsushi Takahashi[2,3,4] *

[1]Biomedical Data Intelligence, Department of Human Health Sciences, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto, 606-8507, Japan. [2]Department of Physical Therapy, School of Health Science and Social Welfare, and [3]Research Institute of Health and Welfare, Kibi International University, Takahashi, Okayama 716-8508, Japan. [4]These authors contributed equally: Yasuhiko Kamikubo, Atsushi Takahashi.

*e-mail: kamikubo.yasuhiko.7u@kyoto-u.ac.jp; atakah7@kiui.ac.jp

Abstract

The outbreak of SARS-CoV-2 in Wuhan, China caused a pandemic of COVID-19. However, it remains enigmatic why the mortality rate is variable among countries. Here we show that at least three types of SARS-CoV-2 virus, type S, K, and G. have spread globally and formed complex infectious trends in terms of transmissibility and virulence. Type K establishes herd immunity and protects against the most virulent type G. Immunity to type S is involved in aggravating type G infections through antibody-dependent enhancement (ADE). Epidemiological tools based on influenza and SARS-CoV-2 epidemic curves explain why COVID-19 mortality varies among Japan prefectures and European countries, and warns of high fatality in the United States. An equation was developed to quantify the severity of COVID-19. Our tools and equations also detect new infectious disease explosions and bioterrorism early, and guide containment of the virus with therapeutic approaches and local policies efficiently inducing herd immunity.

The pandemic of SARS coronavirus 2 (SARS-CoV-2)[1-3] fills the world with fear and confusion, threatening medical collapse and the global financial crisis. Flattening the epidemic curve to avoid loss of healthcare capacity is a major global strategy, but has the disadvantage of delaying the achievement of full herd immunity and sacrificing economic activity. Once the herd immunization is established, the government can confidently decide when to lift restrictions. Therefore, governments need to predict trends in SARS-CoV-2 outbreaks. In a pandemic there is no time to wait for rigorous experimental confirmation. At such times, epidemiology exists first and foremost in determining public health policy. Unfortunately, no government knows the epidemiological methods for predicting their own SARS-CoV-2 status. Here, we develop epidemiological tools to analyze the SARS-CoV-2 epidemic, clarify the trends of the epidemic, and propose countermeasures identified from the analysis.

Sharp drops in the influenza epidemic curve

Japanese doctors have noticed that the flu epidemic has suddenly diminished (*kakeru* in Japanese) this winter season.[4] Type I interferons induced by SARS-CoV-2 may interfere with influenza virus infection.[5] Therefore, an ecological study[6] was conducted to compare the prevalence of SARS-CoV-2 infection with the sum of sentinel influenza surveillance alerts and warnings reported by the Japanese prefectural health centers. The flu may not spread in the summer due to the interference of infectious diseases that spread in the summer, tentatively named *E* (for enteric virus). In tropical and subtropical areas such as Okinawa Prefecture, tropical microorganisms, tentatively called *T*, interferes with *E* in summer, but influenza may not be affected by *T*. This explains why

the flu is prevalent in summer in the tropical and subtropical areas. Okinawa has been excluded from the current analysis due to these potential confounders.

Traces of virus interference starting on January 13, 2020 have been identified in all prefectures of Japan (Fig. 1a). We also noticed a small deflection before the big wave, indicating the possibility of a small pre-epidemic in some areas before SARS-CoV-2 outbreaks (Fig. 1b). Thus, the epidemic curves were stratified into two groups based on the presence or absence of the deflection. One group had a small notch (Fig. 1b) and the other does not (Fig. 1c). The virus that caused the pre-epidemic was tentatively named type S (for *sakigake*, see below), and the virus that caused this major epidemic was tentatively named type K (for *kakeru*).

There was a small spring outbreak (*sakigake* in Japanese) before the Spanish flu. Patients immunized with influenza at that time did not become infected in a subsequent pandemic.[7] Therefore, the prevalence of SARS-CoV-2 was compared with the epidemic curve pattern of each prefecture. Due to the small number of tests in Japan, the prevalence calculated based on the number of patients is not robust. We decided to use the positive rate of the RT-PCR test as the value that best reflects the prevalence. Because the Poisson distribution of small sample results confused the analysis, the study was limited to prefectures with more than 200 tests. Prefectures with the small deflection had significantly lower case fatality rate (CFR) than those without the notch ($P = 0.009$) (Fig. 1d). This suggests that pre-epidemic spread of type S virus induced immunity that reduced the severity of SARS-CoV-2 type K. Therefore, we stratified and analyzed prefectures according to the presence or absence of small deflection.

Analyses of kindergartens in Hokkaido

Hokkaido is the largest of all prefectures in Japan and is divided into many provinces. If these regions are analyzed collectively, effective analysis may not be possible due to the regional heterogeneity. Mathematical modeling has shown that undocumented infections of SARS-CoV-2 were the infection source for most documented cases.[8] Because children infected with SARS-CoV-2 are asymptomatic or mildly symptomatic,[9] it is assumed that children are the primary source of the SARS-CoV-2 outbreak. The incidence of influenza-like illness in kindergartens as reported by health centers in Hokkaido was compared to current SARS-CoV-2 infection. As shown in Fig. 2, strong negative correlation was found (Spearman correlation coefficient ρ= -0.73), suggesting that undocumented SARS-CoV-2 infection among small children in Hokkaido suppressed influenza-like illness. Influenza-like illnesses in older children did not reflect the SARS-CoV-2 epidemic (ρ= -0.14 in elementary school, 0.07 in junior high school, and -0.04 in high school). The flu-like symptoms of COVID-19 in these age groups might have confused the analyses.

Epidemiological description of SARS-CoV-2 epidemics in Japan

The SARS-CoV-2 infection that had subsided in Japan has re-exploded. Only viruses with a higher basic replication number (R_0) can spread to populations with established herd immunity. Presumably, the new virus, derived from type K and carrying a phenotypic mutation, won the competition with the original type K in Europe and the United States, [10] and appears to have entered Tokyo on March 5, 2020. The mutant virus with increased R_0 was tentatively named type G (for global).

In Aichi Prefecture of Japan, patients returning from Hawaii caused an outbreak at a gym. This virus is thought to have been of type G from the United States.

The fatality rate is much higher in Aichi (CFR 8.9%) than other prefectures in Japan (see Fig. 3b). This suggests that type G virus is more virulent than type K. In areas where type G invasion is expected, such as in Aichi, there is a large delayed peak in the SARS-CoV-2 RT-PCR positive rate trend curve (Extended Data Fig. 1a, arrows). On the other hand, this peak is not found in areas where type G has not penetrated, such as Tokyo (Extended Data Fig. 1b). In some areas, the type K peaks (Extended Data Fig. 1c, asterisks) are large and the subsequent type G epidemics are suppressed. We painted the prefectures according to the presence of the K and G peaks (Extended Data Fig. 1d). Type G epidemics will not spread further in prefectures (yellow-green) where both type K and G peaks. On the other hand, in areas where only type K peaks and type G do not reach peak (magenta), it is likely that type G epidemics will progress in the future. In areas where there is no type K peak (red), type G epidemics are expected to be the most intense. This geographical distribution is similar to that of the new SARS-CoV-2 positive cases (Extended Data Fig. 1e), suggesting that herd immunity to the A, K, and G viruses determines how the infection spreads.

Once the herd immunity to type S is established, the proportion (p_S) of the immunized population will be $1 - 1/R_0^S$.[6] When x, y, and z proportions of the population are exposed to type S, K, and G, respectively, the prevalence of infection after the exposure to type K and G (I_K and I_G) is:

$I_K = p_K\, y - p_S\, x = (1 - 1/R_0^K)\, y - (1 - 1/R_0^S)\, x$

$I_G = p_G\, z - p_K\, y = (1 - 1/R_0^G)\, z - (1 - 1/R_0^K)\, y$

Since the prevalence is proportional to the increase in PCR positive rate ([PCR]) when establishing herd immunity,

$I_K = \alpha\, [\mathrm{PCR}]_K$

$I_G = \alpha [\mathrm{PCR}]_G$

By substituting various epidemiological parameters, the constants in these equations can be roughly determined. In Hiroshima, where a clear type S wave is seen (Fig. 1b) and many Chinese tourists visit, the entire population is expected to be exposed to type S and K (x = 1, y =1) and [PCR]w = 0.0018 (0.18%). The highest PCR positive rate has been reported from the Philippines ($[\mathrm{PCR}]_G = 0.573$) as of 3 April 2020. There would have been no previous exposure to either type K or G (x = 0, y =0), and full herd immunity to type G is expected (z = 1). For Hiroshima

$I_K = 1/ R_0^S - 1/ R_0^K = 0.0018\,\alpha$

In the Philippines,

$I_G = 1 - 1/ R_0^G = 0.573\alpha$

The PCR positive rate in Hokkaido rose from 2.17% to 20.83% after the type G virus epidemic (Extended Data Fig. 1a). Thus, $[\mathrm{PCR}]_G = 0.2083 - 0.0217 = 0.1866$, with herd immunity to K and G viruses expected (y= 1, z =1):

$I_K = p_K\,y - p_S\,x = 1 - 1/ R_0^K - (1 - 1/ R_0^S)\,x = 0.0217\alpha$

$I_G = 1/ R_0^K - 1/ R_0^G = 0.1866\alpha$

Solving these equations using $R_0^K = 2.2$[11] gives:

$\alpha = 1.41$, $x_{\mathrm{Hokkaido}} = 0.948$, $R_0^S = 2.19$, $R_0^G = 5.23$. From these calculations, the basic production rate of the new type G virus is considerably higher than the type K virus, meaning a greater increase in transmissibility than the increase in R_0 from type S to K. Nevertheless, it is worth noting that at this R_0 the likelihood of airborne transmission is low.

The difference in PCR positive rates between Japan prefectures may be due to differences in exposure to type S virus. In Aichi (y= 1), RT-PCR positive rates rose from 0.94% (13 February 2020) to 10.06% (6 April), thus $[PCR]_K = 0.0094$, $[PCR]_G = 0.1006$, thus:

$I_K = (1 - 1/ R_0^K)\, y - (1 - 1/ R_0^S)\, x$

$= 0.545y - 0.543x = 0.0094\alpha$

$I_G = p_G\, z - p_K\, y$

$= (1 - 1/ R_0^G)\, z - (1 - 1/ R_0^K)\, y$

$= 0.809z - 0.545y = 0.0966\alpha$

Solving this equation gives:

$x_{Aichi} = 0.980$, $z_{Aichi} = 0.754$

This means that 98.0% and 75.4% of the population in Aichi were exposed to the type S and G viruses, respectively.

In Fukuoka (y= 1), $[PCR]_K = 0.0051$ (16 March 2020), thus:

$I_K = 0.545 - 0.543\, x = 0.0051\alpha$

Solving this equation gives:

$x_{Fukuoka} = 0.991$

In Fukuoka, therefore, 99.1% of the population were exposed to type S SARS-CoV-2. Thus, the proportion of the virus type exposed differs in each region.

When the entire population (y = 1) is exposed to type K or G viruses without viral immunization (x = 0 or y = 0),

$I_K = p_K = 0.545$

$I_G = p_G = 0.809$

These indicate that 54.5% and 80.9% can become infected and cause severe outbreaks. Thus, the outbreak of SARS-CoV-2 in Japan was modest as the spread of type S before type K provided partial immunity to type K.

Analysis of subcluster

According to the Ministry of Health, Labor and Welfare, as of March 29, 2020, 640 out of 1,647 RT-PCR- positive people in Japan are foreigners. Due to the small number of tests in Japan, some ingenuity is needed to understand trends across the population. A 2 x 2 table was created and analyzed according to epidemiological practice. We estimated the number of affected individuals when the proportion of foreigners to Japanese is the same across the population:

c/a = 640/1007= 0.636

	PCR (+)	PCR (-)	
Japanese	a	b	a + b
Foreign	c	d	c + d
	a + c	b + d	

The positive rate of RT-PCR in the general population on Japan on March 26 is as follows:

a/(a + b)= 0.0624 (1647/26401)

The approximate total population of Japanese is:

$a + b = 126.8 \times 10^6$

Assuming that the proportion of foreigners in the total population is f,

$f = (c + d)/(a + b)$

Solving these equations gives:

$a = 7.91 \times 10^6$

$b = 118.89 \times 10^6$

$c = 5.03 \times 10^6$

$d = 7.91 \times (16.03f — 0.636) \times 10^6$

$d > 0$, so $f > 0.0397$ and $c + d > 5.03 \times 10^6$, imposing constraints on this variable. The constraint indicates that more than 5.03×10^4 foreigners must be present in Japan. This estimate is plausible, as there were 2.67 million foreigners registered in Japan in early 2020, and 1.84 million Chinese tourists came to Japan from the end of last year to the beginning of this year. Most of the infected 5.03×10^6 foreigners may have returned to their counties. If $f = 0.04$,

$d = 0.589 \times 10^6$

The prevalence of SARS-CoV-2 among foreigners is as follows:

$c/(c + d) = 0.749$ (74.9%) ($P = 0$, McNemar's test)

If this subpopulation has established herd immunity,

$R_0 = 1/(1- 0.749) = 3.99$

This suggests that foreigners had formed highly contagious subclusters. Such analyses can help determine the size and density of the dense contact subcluster on epidemics.

Development of risk scores

Virus-to-virus interference occurs because the biological response to one virus prevents the transmission of the next virus.[5] Therefore, the extent to which SARS-CoV-2 affects the influenza epidemic curve should be determined by the strength of the biological

response. The small wave of type S infection during the week from December 23 (Fig. 1b) is not necessarily due to the small number of people infected. This may have been due to weak biological response to type S infection. Response to type S may have been weaker due to previous exposure to a more ancestral virus or due to competing viruses. The asymptomatic rate of Japanese evacuated from Wuhan, which is highly likely to be infected with type K and/or G, is 30.8% (95% CI: 7.7-53.8%).[12] In contrast, passengers on the cruise ship Diamond Princess were considered to be type S and/or K because they were tracked to Hong Kong passengers, and 51.4% (318/619) of RT-PCR confirmed cases were asymptomatic.[13] Due to the high asymptomatic proportion of cruise ship passengers, type S infection is more likely to be asymptomatic than type K or G, supporting the view that the virus can easily escape surveillance recognition.

An immune response to the K virus occurred following the K infection found in all prefectures in the week beginning January 13 (Fig. 1a). Cytokine storms are implicated in severe COVID-19.[14] The strong immune response can also affect the flu epidemic, causing a decline in the epidemic curve. A weak response to SARS-CoV-2 may not be enough to interfere with the flu. In fact, a small number of cases have been reported to be co-infected with SARS-CoV-2 and influenza virus.[15] The epidemic curves after type K wave differed by prefecture and showed clearly different degrees of decline (Extended Data Fig. 2). Therefore, we have developed a scoring system for epidemic curve reduction as a surrogate indicator of innate and acquired immunity to type K virus.

The degree of immunity to SARS-CoV-2 types S, K, and G determines the prevalence of infection, as described above. Therefore, the degree of herd immunity can be estimated from the positive rate of RT-PCR. However, performing RT-PCR tests on

many patients places a heavy burden on society and healthcare professionals, and there has been a growing need for easily available surrogate indicators. Indeed, our proxy for immunity to SARS-CoV-2, now termed the "risk score", was positively correlated with the prevalence of SARS-CoV-2 (Spearman correlation coefficient $\rho = 0.415$) (Fig. 4a).

Immunological basis for transition to type G

Because the high mortality rate of type G patients in Aichi and Hokkaido (Fig. 3b) confounds the analysis, stratification was performed to exclude Aichi and Hokkaido from the comparison with CFR. Still, the correlation with severity of infection was weaker than prevalence (Spearman correlation coefficient $p = 0.365$) (Extended Data Fig. 3a).

To determine why the immune response to SARS-CoV-2 is not reflected in mortality, we calculated the effects of type S, K, and G infections on mortality. The reduction in CFR (F) of the present type caused by previous infection by another type was defined as an attenuating factor (a). When calculated using the prevalence (I) and the exposure to type S (x) , K (y), and G (z):

$$F^K = F_0^K y - a_{(K\text{-}S)} I_S$$

$$= F_0^K y - a_{(K\text{-}S)} (1 - 1/ R_0^S) x$$

$$F^G = F_0^G z - a_{(G\text{-}K)} I_K - a_{(G\text{-}S)} I_S$$

$$= F_0^G z - a_{(G\text{-}K)} [(1 - 1/ R_0^K) y - (1 - 1/ R_0^S) x] - a_{(G\text{-}S)} x (1 - 1/ R_0^S)$$

$$= F_0^G z - a_{(G\text{-}K)} (1 - 1/ R_0^K) y + (a_{(G\text{-}K)} - a_{(G\text{-}S)}) (1 - 1/ R_0^S) x$$

Again, we roughly determined the constants in these equations by substituting epidemiological parameters. In the Philippines (x = 0, y = 0, z = 1):

$F^G = F_0^G z - a_{(G\text{-}K)} I_K - a_{(G\text{-}S)} I_S = F_0^G = 104/3764 = 4.70$ % (7 April 2020)

In Hokkaido (y = 1. z = 1), $I_K = 0.0217\alpha$ and $x_{Hokkaido} = 0.948$ as shown above. thus:

$F^G = F_0^G\, z - a_{(G\text{-}K)}\, I_K - a_{(G\text{-}S)}\,(1 - 1/R_0^S)\, x_{Hokkaido}$

$= 4.70 - 0.0217\alpha\, a_{(G\text{-}K)} - 0.543\, a_{(G\text{-}S)}\, x_{Hokkaido} = 4.19\%$ (7 April 2020)

In Aichi (y = 1), $x_{Aichi} = 0.980$, $z_{Aichi} = 0.754$, $I_K = 0.0094\alpha$, and $I_G = 0.0966\alpha$ (6 April 2020):

$F^G = F_0^G z - a_{(G\text{-}K)}\, I_K - a_{(G\text{-}S)}\,(1 - 1/R_0^S)\, x_{Aichi}$

$= 4.7\, z_{Aichi} - 0.0094\alpha\, a_{(G\text{-}K).} - 0.543\, a_{(G\text{-}S)}\, x_{Aichi} = 8.81\%$

Ehime Prefecture had the highest mortality rate before the appearance of the G wave (Fig. 3b). The foreign tourist visit rate of Ehime is as low as 0.4%, ranking 39th among 47 prefectures in Japan.[17] Moreover, only 12.29% are Chinese. Ehime Prefecture does not have any of the top 30 most popular tourist spots for foreign visitors to Japan.[18] We therefore expect that type S virus failed to enter Ehime (x = 0):

$F^K = F_0^K - a_{(K\text{-}S)}\,(1 - 1/R_0^S)\, x = F_0^K = 4.76\%$

In Fukuoka, $x_{Fukuoka} = 0.991$, y=1, and CFR is 0.62%:

$F^K = F_0^K y - a_{(K\text{-}S)}\, I_S = F_0^K - a_{(K\text{-}S)}\,(1 - 1/R_0^S)\, x_{Fukuoka} = 4.76 - 0.538\, a_{(K\text{-}S)} = 0.62$

These equations have been solved using $\alpha = 1.411$:

$a_{(K\text{-}S)} = 7.69$, $a_{(G\text{-}K)} = 313.43$, $a_{(G\text{-}S)} = -17.66$

$F^K = F_0^K\, y - a_{(K\text{-}S)}\,(1 - 1/R_0^S)\, x$

$= 4.76\, y - 4.18\, x$

$F^G = F_0^G z - a_{(G\text{-}K)}\,(1 - 1/R_0^K)\, y + (a_{(G\text{-}K)} - a_{(G\text{-}S)})(1 - 1/R_0^S)\, x$

$= 4.7\, z - 170.96\, y + 179.75\, x$

Interestingly, $a_{(G\text{-}S)} < 0$, indicating that preexisting type S infection enhances rather than attenuates the severity of the closest type G infection. Antibody-dependent enhancement (ADE) has been implicated in SARS becoming virulent.[19] ADE is a phenomenon in

which antibodies promote virus entry into cells through Fcγ receptors (see Fig. 7c), and are also found in dengue, yellow fever,[20] and human immunodeficiency viruses. ADE is also thought to cause a sharp rise in R_0, which is consistent with the type G properties. Total CFR in the population is:

$F = F^K + F^G$

$= F_0^G z - [a_{(G\text{-}K)}(1 - 1/R_0^K) - F_0^K]y + (a_{(G\text{-}K)} - a_{(G\text{-}S)})(1 - 1/R_0^S)x$

$= 4.7\,z - 166.20\,y + 175.58\,x$

Thus, the equation, which we playfully call the Kami (for Kamikubo) -Atsushi (sounds like “God is merciful” in Japanese) equation, explains how SARS-CoV-2, a virus of common cold（*kaze* in Japanese）, has become a virulent pneumonia virus.[1-3] Considering the contribution of each variable, it can be seen that ADE caused by S virus is predominantly involved in the fatal condition. Depending on the rate at which the population is exposed to each type of virus (S: x, K: y, G: z), the severity and mortality when exposed to type K (4.76 y – 4.18 x) and type G (4.7 z – 170.96 y + 179.75 x) can vary widely. In particular, the mortality of type G virus will jump up (4.7 z + 179.75 x) if the population were not exposed to type K after exposure to type S.

The risk score for the magnitude of the inflammatory response inferred from the influenza epidemic curve does not incorporate the benefits of immunity to type K and the negative side of type S ADE. Therefore, we estimated K and G infections from the SARS-CoV-2 epidemic curve and quantified it as a scoring system that reflects the proportion of G infections in the epidemic, ie., z / (y + z). The score, named “G-score”, correlated strongly with CFR (Fig. 3b), supporting one of the predictions of Kami-Atsushi equation that the negative effects of immune responses from type G infection determine the severity of the SARS-CoV-2 epidemic when $F_0^G > 0$. The scores for each

prefecture are displayed on a map of Japan (Fig. 3c). Prefectures with high G scores will have a higher CFR due to the higher percentage of type G COVID-19 than type K. Areas with low G scores indicate a low degree of herd immunity to type G, and are expected to increase prevalence during type G invasion.

Genetic basis for transition to type G

The antibodies that cause ADE are directed against the viral spike protein. Higher antibody binding neutralizes virus entry, while lower binding results in ADE.[21] This means that mutations in the spike protein of SARS-CoV-2 weaken the binding of herd immunity antibodies and induce ADE. We therefore examined the phylogenic distribution of mutations in the spike protein. As shown in Fig. 4a, the S614G spike mutation segregated a group, which is likely to be the type G, and was found to be under high positive selection pressure. This amino acid is contained in the region identified as a dominant B cell epitope.[22] The S614G spike mutation occurred independently in Wuhan (Fig. 4b), suggesting convergent evolution. This mutation probably occurred in Shanghai (Fig. 4a). The most likely story is that the mutant virus flew to Italy, where it acquired the P314L mutation in the *ORF1b* gene (Fig. 4c), and then spread to Europe, the United States, and later to Asia (Fig. 4d). This paralleled the news of higher mortality rates, suggesting that ADE is also associated with the heightened pathogenicity.[23] European physicians have reported cases in which the condition worsened rapidly despite a decrease in viral load.[24] Because ADE can cause the virus to enter cells quickly, the amount of virus detected decreases rapidly. Also, asymptomatic patients may worsen rapidly because ADEs suddenly occur when patients begin to produce low-affinity antibodies to the spike protein. Recently, a fatal illness

was reported at age 0,[9] which also suggests maternal antibody ADE. These results indicate that ADE's involvement in COVID-19 illustrates the current state of the world pandemic.

Type G is a variant of the virus reported as “L type” in population genetic studies (Extended Data Fig. 4).[25] Further research is needed on where the type S and K belong in the phylogenetic tree. However, the virus reported as “S type” is more ancestral and appears to be the most appropriate candidate for type S in the present study.

Next, we analyzed epidemiological trends in China. In Wuhan, CFR was 2.9%:[26]

$F = = 4.7\ z - 166.20\ y + 175.58\ x = 2.9$

Assuming that L type is mixture of type K and G, the frequency of L type and S type was 96.3% and 3.7% in Wuhan, respectively:[25]

$y + z = 0.963$

$x = 0.037$

Assuming that three kinds of viruses were simultaneously epidemic, and residents were exposed to any of the viruses:

$x + y + z = 1.0$

Solving the equation gives:

$y = 0.0475, z = 0.9155.$

Outside Wuhan, CFR was 0.4%[26] and the frequency of L and S types was 61.6% and 38.4%, respectively:[25]

$F = 4.7\ z - 166.20\ y + 175.58\ x = 0.4$

$y + z = 0.616$

$x = 0.384$

Solving the equation gives:

$y = 0.409$, $z = 0.207$.

This result indicates that most of the viruses that spread in Wuhan were type G, while outbreaks outside of Wuhan were more caused by type K than by type G. Thus, differences in mortality can be explained without considering medical collapse in Wuhan as a risk factor.

Why mortality rates differ in European countries

Europe has become the center of the pandemic. However, why mortality rates vary from country to country remains enigmatic.[27] Influenza epidemic curves in European countries varied greatly from one country to another (Extended Data Figs. 5 and 6). Risk scores were calculated for European countries. This score correlated with SARS-CoV-2 CFR in each country (Fig. 5a). The risk score essentially reflects $\beta = (y + z) / y$:

$F = 4.7\ z - 166.20\ y + 175.58\ x$

$= 4.7\ (\beta - 167.20)\ y + 175.58\ x$

When $x \gg y$ or z as observed in Japan, CFR (F) does not closely correlate with the risk score (β). In Europe, however, x seems to be sufficiently low so that the correlation is clearly discerned. When $y = 1$ and $x \sim 0$,

$F = 4.7\beta - 167.20$

The results suggest that severity and mortality in Europe are determined by herd immunity to the type K virus. Displaying the risk score on a map reveals geographical effects on the establishment of herd immunity (Fig. 5b).

Because the SARS-CoV-2 type S waves in the influenza epidemic curve are weak and difficult to see, the notches that are clearly visible in many European countries (Extended Data Figs. 5 and 6) should indicate multiple invasions of type K viruses. In fact, the phylogenetic tree of the viral genome suggests that there are multiple K subtypes. If only small notches are observed, it is likely that only type S has invaded or that the prior spread of type S has suppressed the inflammatory response to type K. The absence of type K waves in the epidemic curves of Italy, Ukraine, Spain, the UK, and France suggests poor type K protective immunity against type G, while the ADE by type S should exacerbated COVID-19.

Predicting COVID-19 severity in the United States

The risk score seems to predict mortality in the United States (Fig. 6a). States with higher scores but lower CFR may have lower total virus tests. The risk score also predicts that there are many areas where protective herd immunity to SARS-CoV-2 type K is not expected (Fig. 6b). The outbreak of influenza this winter clearly indicates that there was little spread of SARS-CoV-2 type K in the United States. Like the deadliest countries in Europe, many states in the United States have poorly shaped K-waves (Extended Data Figs. 7 and 8). In these regions, not only does protective immunity by K virus fail, but ADE due to immunity to type S can contribute to aggravation. Compared to the maps of Japan and Europe (Extended Data Fig. 3a, Fig. 3c, and Fig. 5b), this is a warning result suggesting that many states in the United States will have a higher COVID-19 mortality than Japan or Europe.

Discussion

Our epidemiological analysis indicates that a virulent type G SARS-CoV-2 has R_0 as high as 6.99. If type G virus spreads to populations that have not been previously exposed to the type S or K, it will lead to a relatively fatal disease with a prevalence up to 85.7%. In addition, ADE is thought to occur in type G infection, and patients producing low affinity antibodies are expected to experience exacerbations of COVID-19[23] such as cytokine storms. Healthcare professionals in the West must keep in mind that the prevailing virus is of a nature beyond experiences in East Asia. It is important to understand the paradoxical situation[28] that deteriorates when a subgroup exposed to type S and subsequently protected from type K (Fig. 7a) is exposed to type G (Fig. 7c). In the future, if a new mutant virus attacks, it is assumed that such reversal phenomenon will occur one after another, and caution is required. This means that there are additional spike protein mutations that can cause ADE, but the number of such epitopes should be limited.[22]

Japan's low mortality rate is due to the influx of type K throughout the country in the week of January 13 and the establishment of herd immunity that confers resistance to type G. Until 9 March 2020, the Japanese government did not restrict entry of tourists from areas of China outside Hubei Province, unintentionally permitting the selective influx of type S and K into Japan by 1.84 million Chinese tourists. The return of Japanese refugees on a charter aircraft from Wuhan resulted in the influx of the K virus, which led to wider herd immunity and readiness for the next type G outbreak. In the United States and most European countries, travel from China was restricted from the beginning of February, blocking type K influx from China. There is deep concern that the serious outbreak of SARS-CoV-2 in the United States and Europe may have been due to the spread of type G before adequate immunity to types S and K was established. Future

epidemic control measures must consider options to establish herd immunity before more virulent viruses enter by allowing for the influx of attenuated viruses.

We have developed an epidemiological tool to monitor and predict COVID-19 severity based on influenza epidemic curves (Extended Data Fig. 9). This tool can help balance social isolation and herd immunity policies and guide the development of public health policies that are effective and minimize economic losses. COVID-19 might be more severe than Wuhan, China in areas where herd immunity has not been established by influx of type K.

No test makes sense without knowing the true prevalence. To determine the normal range of the test, you need to determine the test values for unaffected individuals. Determining a cut-off value[6] can be difficult for diseases with a high prevalence of subclinical infections, such as SARS-CoV-2, especially when subclinical infections are prevalent in most populations.

Careful monitoring of the influenza epidemic curve (Fig. 1b and Extended Data Fig. 9) can reveal the emergence of microorganisms that interfere with influenza infection, such as SARS-CoV-2. This method can be used universally among competing viruses and other microorganisms. Of course, there are several infections that can interfere with influenza, and there may be co-existing infections that interfere with that infection. The sharp change in the epidemic curve that took place across the country as of January 13 (Fig. 1a) is a strong warning sign of the rapid spread of powerful infections that may interfere with influenza. Public health measures at this stage may have stopped the epidemic. Therefore, it is very useful to use microbial interference to detect the occurrence of a new infection and take immediate action. The Global Early Warning System, a Canadian health monitoring platform based on "BlueDot"[29] using AI, noticed

the outbreak of pneumonia in Wuhan on December 31, 2019. If we were able to notice a small deflection due to the type S virus from December 23, 2019 (Fig. 1b), epidemiological tools would have been able to detect the virus landing in Japan sooner, demonstrated that a better security system could be developed using human thinking instead of blindly relying on AI. It is worth noting that this system is also useful for detecting bioterrorism.

Our epidemiological analyses also provide hints on the diagnosis and treatment of COVID-19. Children and pregnant women, who are more likely to produce low-affinity antibodies, may be at risk for severe COVID-19 from ADE. If the virus enters the cells due to ADE, the measurement of free virus does not reflect the pathophysiology. Technology for measuring intracellular viruses should be developed. If ADE is exacerbating the condition, plasmapheresis or antibody-adsorbing columns may be effective in treating severe cases and preventing deterioration. Convalescent sera may be useful for treatment. However, sera containing antibodies with low affinity for spike proteins can induce ADE and, conversely, lead to more severe disease. If the affinity of the antibody is high, the neutralizing effect becomes dominant, so a high-affinity antibody against the viral spike protein may subside the ADE pathology.[30] Furthermore, gene editing that reverses the S641G mutation of type G virus may be an effective therapeutic strategy. Attenuated viruses and bacteria have been used as pesticides to protect against plant diseases such as cucumber wilt, tomato blight, and sweet potato fusarium wilt.[31] In the absence of an effective vaccine, a strategic infection may need to be considered to overcome the ADE and achieve the required immunity.

Of course, epidemiological research is not without its pitfalls. The presence of confounders can lead to incorrect conclusions about epidemiological studies. The results

of our analysis also require constant scrutiny based on scientific skepticism for any confounding factors that have not yet been noticed. Taken together, however, it is clear that historical lessons indicate that epidemiology should be used during a pandemic.

Methods

Sources of data

Epidemiological data were obtained from the Worldometer,[32] the COVID Tracking Project,[33] the Hokkaido Infectious Disease Surveillance Center,[34] and websites operated by Su Wei,[35] the Ministry of Health, Labor and Welfare, Japan,[36] the Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Japan,[37] the Hokkaido government,[38] and the World Health Organization.[39]

Statistical analysis

The quantification and modeling of the epidemiological curve was performed under contract with Shin Nippon Biomedical Laboratories, Ltd. (Tokyo, Japan). Analyses of Spearman correlation coefficient were performed with the use of the Statcel4 add-in package (OMS Publishing, Tokorozawa, Japan) for Microsoft Excel.

Data availability

All data are available on request.

References

1 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269, doi:10.1038/s41586-020-2008-3 (2020).

2 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273, doi:10.1038/s41586-020-2012-7 (2020).

3 Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727-733, doi:10.1056/NEJMoa2001017 (2020).

4 Sakamoto, H., Ishikane, M. & Ueda, P. Seasonal Influenza Activity During the SARS-CoV-2 Outbreak in Japan. *JAMA*, doi:10.1001/jama.2020.6173 (2020).

5 Murphy, K. & Weaver, C. *Janeway's Immunobiology*. 9th edn, (Garland Science, 2017).

6 Giesecke, J. *Modern Infectious Disease Epidemiology*. 3rd edn, (CRC Press, 2017).

7 Crosby, A. W. *America's Forgotten Pandemic: The Influenza of 1918*. (Cambridge University Press, 2003).

8 Li, R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, doi:10.1126/science.abb3221 (2020).

9 Lu, X. *et al.* SARS-CoV-2 Infection in Children. *N Engl J Med*, doi:10.1056/NEJMc2005073 (2020).

10 GISAID. *Genetic epidemiology of hCoV-19*, <https://www.gisaid.org/epiflu-applications/next-hcov-19-app/> (2020).

11 Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med*, doi:10.1056/NEJMoa2001316

(2020).

12 Nishiura, H. *et al.* Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int J Infect Dis*, doi:10.1016/j.ijid.2020.03.020 (2020).

13 National Institute of Infectious Diseases, J. *Field Briefing: Diamond Princess COVID-19 Cases, 20 Feb Update*, <https://www.niid.go.jp/niid/en/2019-ncov-e/9417-covid-dp-fe-02.html> (

14 Mehta, P. *et al.* COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* **395**, 1033-1034, doi:10.1016/S0140-6736(20)30628-0 (2020).

15 Ding, Q., Lu, P., Fan, Y., Xia, Y. & Liu, M. The clinical characteristics of pneumonia patients coinfected with 2019 novel coronavirus and influenza virus in Wuhan, China. *J Med Virol*, doi:10.1002/jmv.25781 (2020).

16 Wikipedia. *COVID-19 testing*, <https://en.wikipedia.org/wiki/COVID-19_testing> (

17 Honichi Lab. *[Ehime Prefecture Inbound Demand] (Japanese)*, <https://honichi.com/areas/chugokushikoku/ehime/> (

18 Honichi Lab. [Top 30 inbound tourist destinations] (Japanese).

19 Yang, Z. Y. *et al.* Evasion of antibody neutralization in emerging severe acute respiratory syndrome coronaviruses. *Proc Natl Acad Sci U S A* **102**, 797-801, doi:10.1073/pnas.0409065102 (2005).

20 Katzelnick, L. C. *et al.* Antibody-dependent enhancement of severe dengue disease in humans. *Science* **358**, 929-932, doi:10.1126/science.aan6836 (2017).

21 Wang, S. F. *et al.* Antibody-dependent SARS coronavirus infection is mediated by antibodies against spike proteins. *Biochem Biophys Res Commun* **451**, 208-

214, doi:10.1016/j.bbrc.2014.07.090 (2014).

22 Grifoni, A. *et al.* A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe*, doi:10.1016/j.chom.2020.03.002 (2020).

23 Tetro, J. A. Is COVID-19 receiving ADE from other coronaviruses? *Microbes Infect* **22**, 72-73, doi:10.1016/j.micinf.2020.02.006 (2020).

24 Lescure, F. X. *et al.* Clinical and virological data of the first cases of COVID-19 in Europe: a case series. *Lancet Infect Dis*, doi:10.1016/S1473-3099(20)30200-0 (2020).

25 Tang, X. *et al.* On the origin and continuing evolution of SARS-CoV-2. *National Science Review*, doi:10.1093/nsr/nwaa036 (2020).

26 Novel Coronavirus Pneumonia Emergency Response Epidemiology, T. [The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China]. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 145-151, doi:10.3760/cma.j.issn.0254-6450.2020.02.003 (2020).

27 Onder, G., Rezza, G. & Brusaferro, S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA*, doi:10.1001/jama.2020.4683 (2020).

28 Morita, K. *et al.* Paradoxical enhancement of leukemogenesis in acute myeloid leukemia with moderately attenuated RUNX1 expressions. *Blood Adv* **1**, 1440-1451, doi:10.1182/bloodadvances.2017007591 (2017).

29 BlueDot. *Global Early Warning System*, <https://bluedot.global/products/> (

30 Shen, C. *et al.* Treatment of 5 Critically Ill Patients With COVID-19 With Convalescent Plasma. *JAMA*, doi:10.1001/jama.2020.4783 (2020).

31 Gottula, J. & Fuchs, M. Toward a quarter century of pathogen-derived resistance and practical approaches to plant virus disease control. *Adv Virus Res* **75**, 161-183, doi:10.1016/S0065-3527(09)07505-8 (2009).

32 Worldmeter. *Worldometerhttps://www.worldometers.info/*, <https://www.worldometers.info/> (

33 Project, T. C. T. *Most recent data*, <https://covidtracking.com/data/> (

34 Center, H. I. D. S. *[Pathogen / School Health / Inpatient Information] (Japanese)*, <http://www.iph.pref.hokkaido.jp/kansen/otherscomment.html> (

35 Wei, S. *[Bulletin of SARS-CoV-2 infection] (Japanese)*, <https://covid-2019.live/> (

36 Ministry of Health, L. a. W., Japan. *About Coronavirus Disease 2019 (COVID-19)*, <https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/newpage_00032.html> (

37 Infectious Disease Surveillance Center, N. I. o. I. D., Japan. *[Influenza epidemic level map] (Japanese)*, <https://nesid4g.mhlw.go.jp/Hasseidoko/Levelmap/flu/2019_2020/trend.html> (

38 government, H. *[Coronavirus Disease 2019: Outbreak in Hokkaido](Japanese)*, <http://www.pref.hokkaido.lg.jp/hf/kth/kak/hasseijoukyou.htm> (

39 Europe, F. N. Comparison of number of influenza detections by subtype.

Acknowledgements We thank H. Asakura for suggesting epidemiological analyses of COVID-19, E. Ogawa for collecting samples, H. Kitano for suggesting viral competition, T. Hattori for suggesting ADE, S. Mayama for pointing out attenuated viruses used as pesticides, K. Kitagawa, K. Matsushima, and M. Naruke for discussions. This work was supported by Grant-in-Aid for Scientific Research (KAKENHI; 17H03597 and

16K14632) from the Japan Society for the Promotion of Science.

Author contributions Y.K. and A.T. conceived and designed the study. Y.K. performed the sample collection. A.T. performed the epidemiological and statistical analyses. Y.K. and A.T. wrote the paper.

Competing interests The authors declare no competing interests.

Additional Information

Supplementary information is available for this paper at https://

Correspondence and requests for materials should be addressed to Y.K. and A.T.

Figure legends

Fig. 1 | SARS-CoV-2 trend curves in Japan prefectures. a, The numbers of influenza alerts and warnings reports from November 11, 2019 to March 8, 2020 from all prefectures in Japan have been plotted. **b,** Trend curves of prefectures with small deflections before large waves. **c,** Prefectural trend curves without small deflections before large waves. **d,** Box-and-whisker plot comparing the prevalence of RT-PCR in prefectures with (+) and without (-) small deviations before large waves.

Fig. 2 | Correlation between influenza in kindergartens and SARS-CoV-2 infection. The correlation diagram shows the prevalence (%) of influenza-like illness in kindergartens in each district in Hokkaido and the number of individuals who were RT-PCR positive for SARS-CoV-2. Spearman correlation coefficient $\rho = -0.73$.

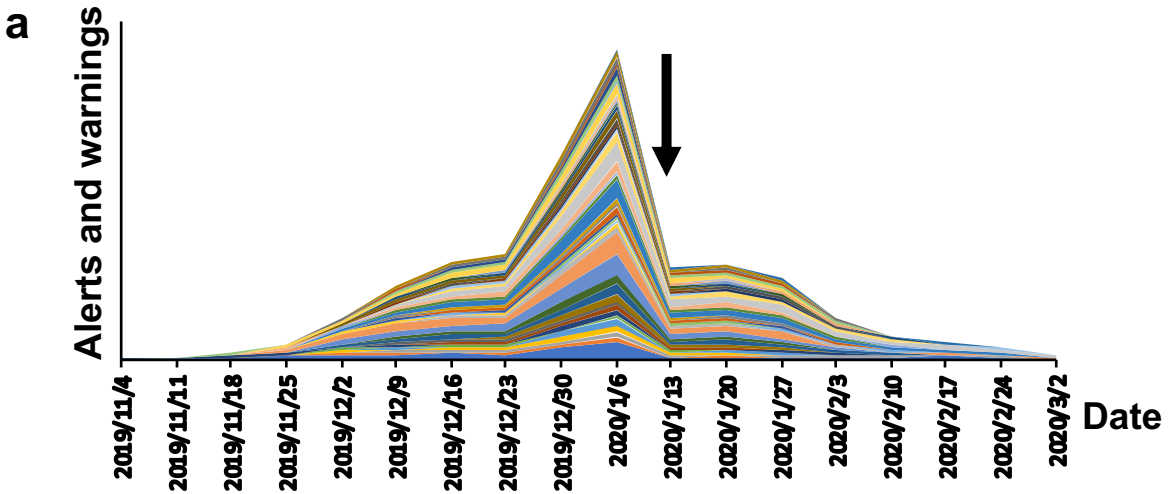
Fig. 3 | The risk score, G score, and COVID-19 mortality in Japan. a, Correlation between the risk score and RT-PCR positive rates (%) for SARS-CoV-2 in Japan prefectures. Spearman correlation coefficient $\rho = 0.415$. **b,** Correlation between the G score and CFR of COVID-19 in Japan prefectures. Spearman correlation coefficient $\rho = 0.537$. **c,** Japan map showing the distribution of G scores of various strengths between prefectures.

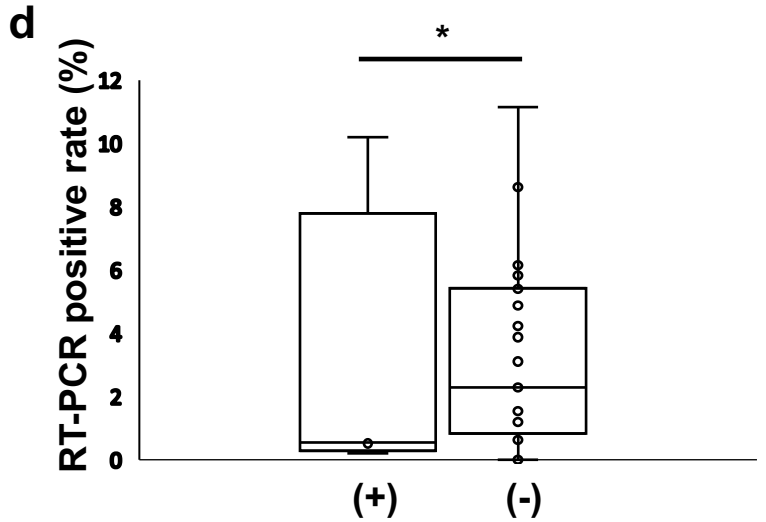
Fig. 4 | Genomic analysis of type G SARS-CoV-2. a, S614G mutation in the SARS-CoV-2 spike protein in Shanghai. **b,** Independent S614G mutation in Wuhan. **c,** *ORF1b* P314L mutation in Italy. **d,** Outbreak the type G virus in Europe and the United States and further spread to Asia.

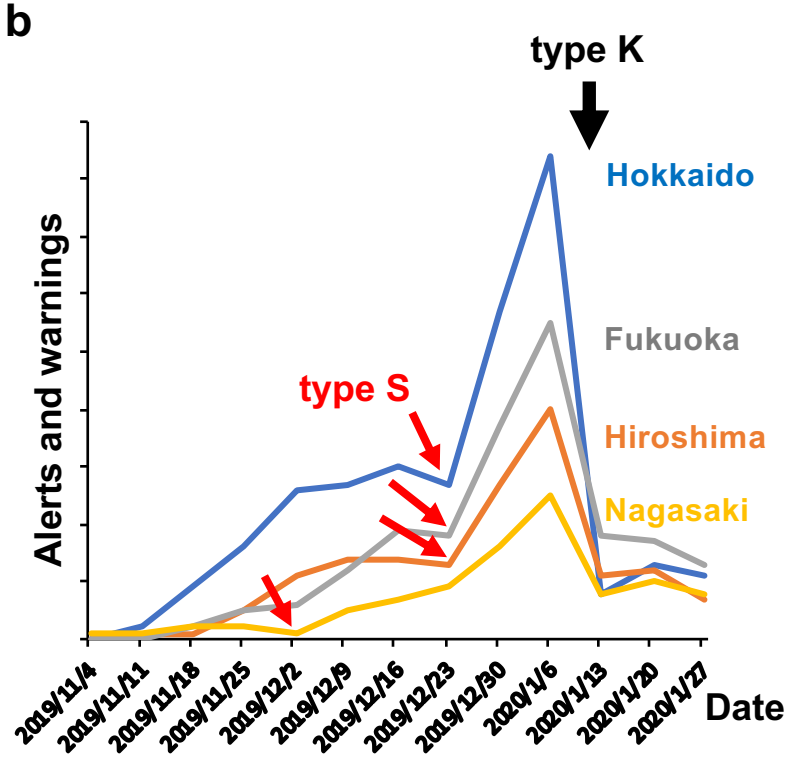
Fig. 5| The risk scores and COVID-19 mortality in Europe. a, Correlation between the risk score and CFR of COVID-19 in European countries. Spearman correlation coefficient ρ = 0.65. **b,** A map of Europe showing the distribution of risk scores for different strengths. Influenza epidemic data was not available in the blank area.

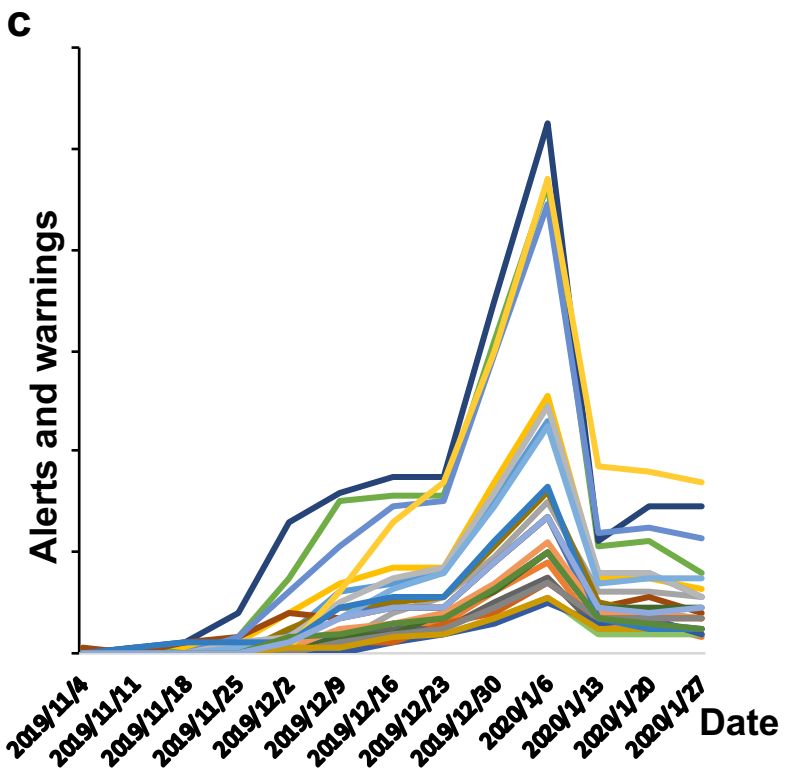
Fig. 6 | Risk scores and CFR in USA. a, Correlation between Risk Scores and CFR of COVID-19 in U.S. States. Spearman correlation coefficient ρ= 0.46. **b,** European map.

Fig. 7 | A model for antibody-dependent enhancement of SARS-CoV-2 type G infection. a, By specifically binding to type G SARS-CoV-2 surface spikes, neutralizing antibodies prevent interaction with host cells that can be infected and destroyed. **b,** Infection or COVID-19 in the absence of antibodies. **c,** The enhancing antibody helps type G SARS-CoV-2 infect monocytes more efficiently. It increases the overall replication of the virus and the risk of severe type G COVID-19. Thus, antibody-dependent enhancement (ADE) increases the infection of SARS-CoV-2 particles during subsequent COVID-19 by another viral serotype.

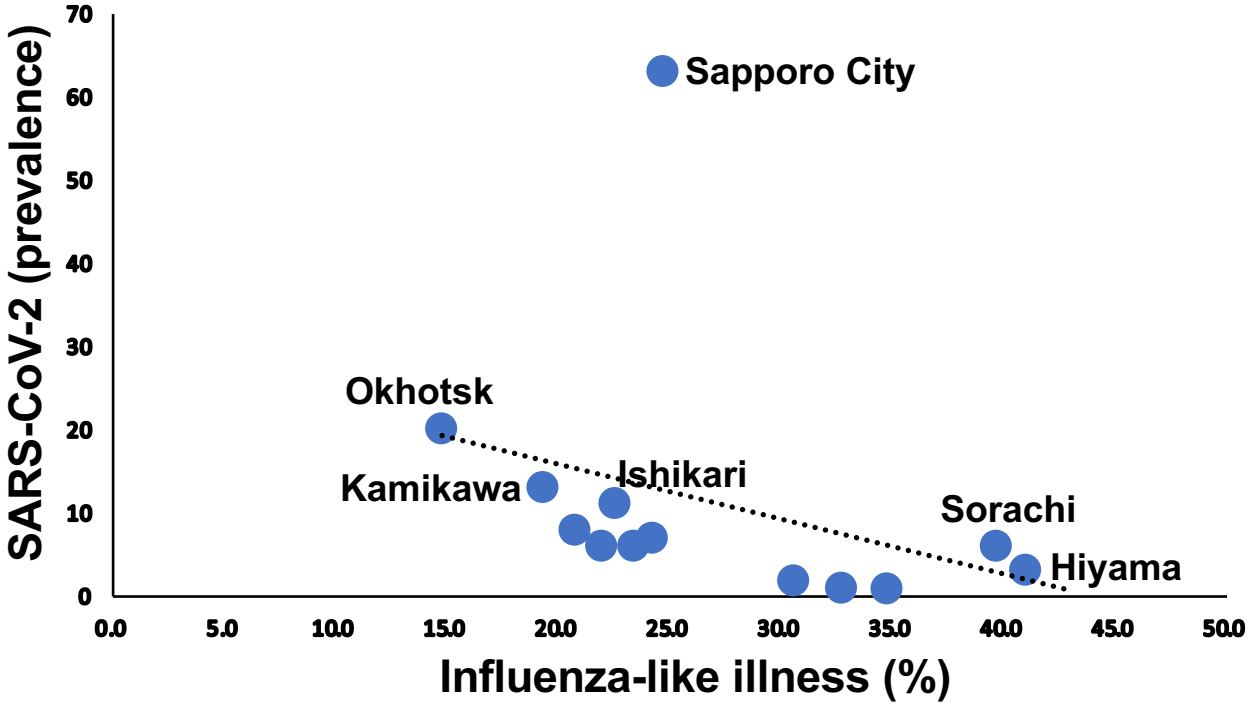




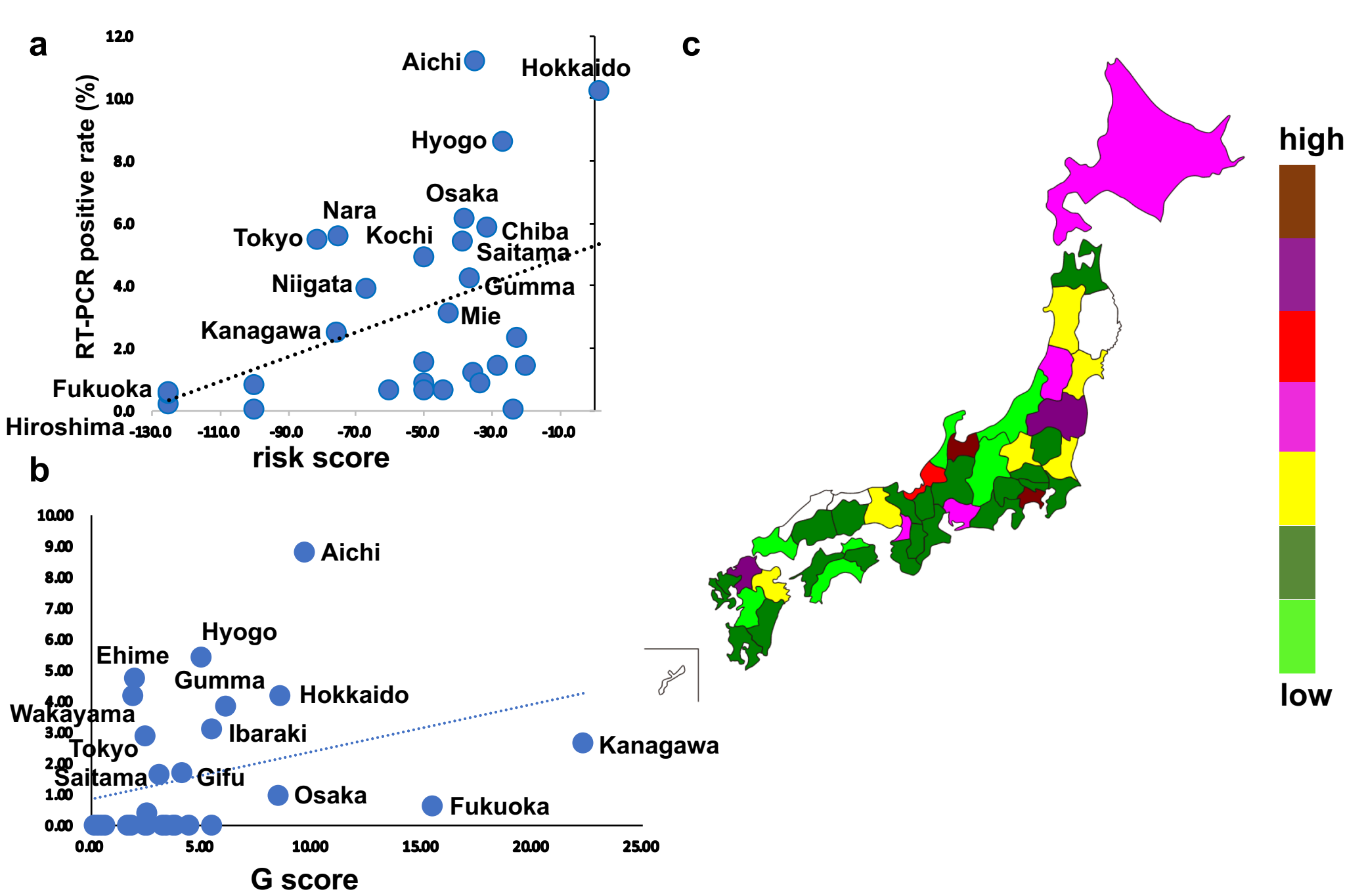




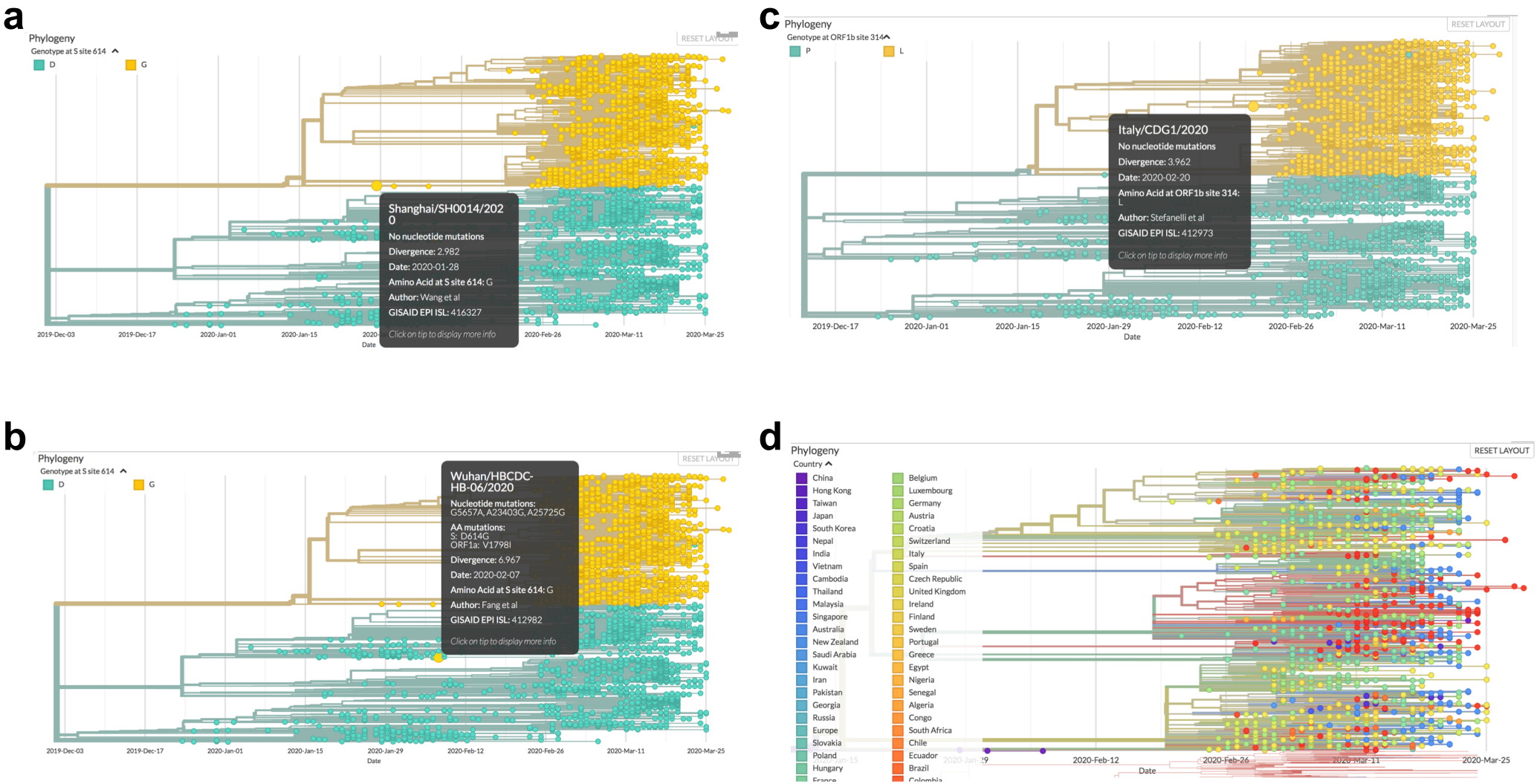
Kamikubo & Takahashi, Fig. 1



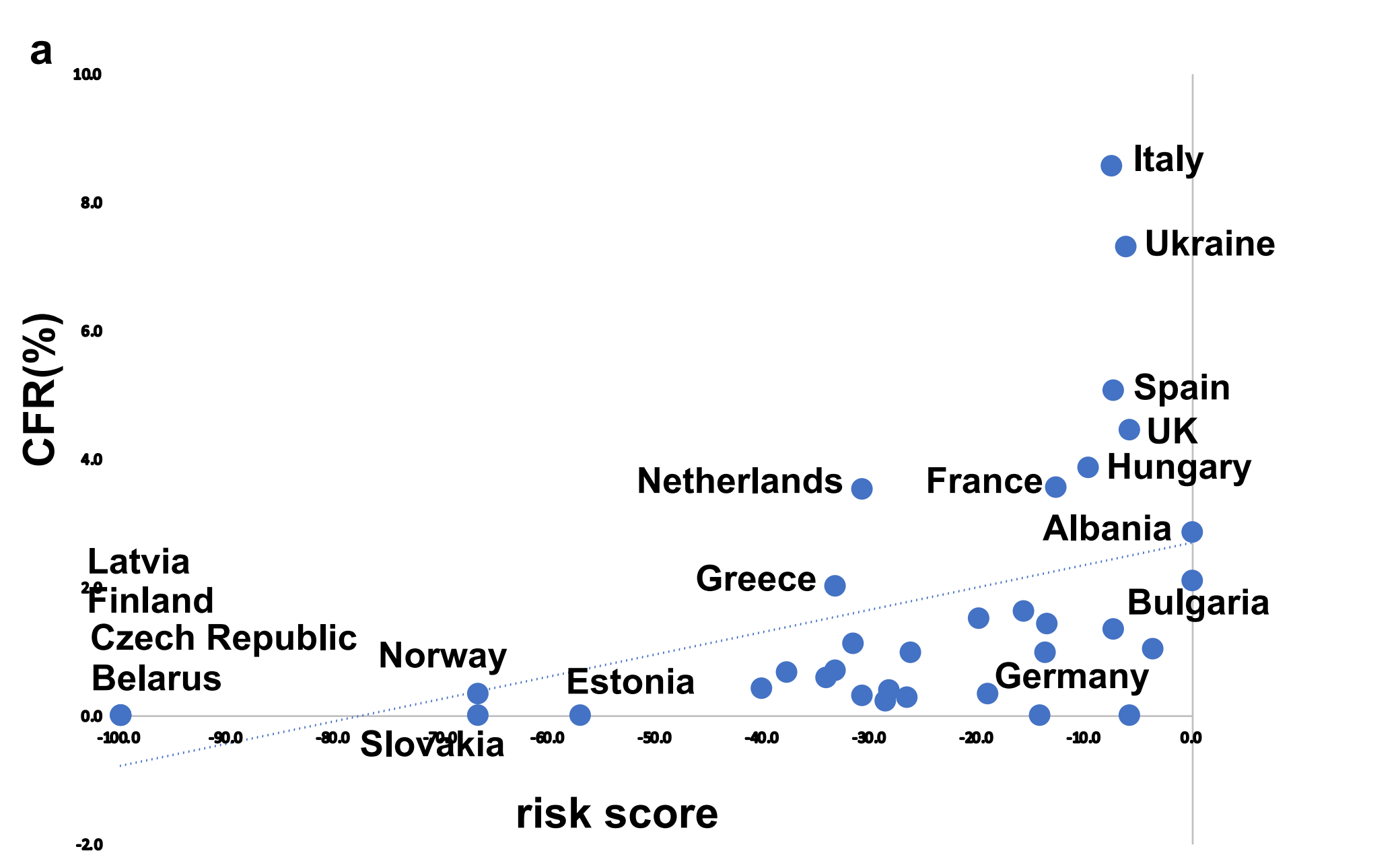
Kamikubo & Takahashi, Fig. 2



Kamikubo & Takahashi, Fig. 3

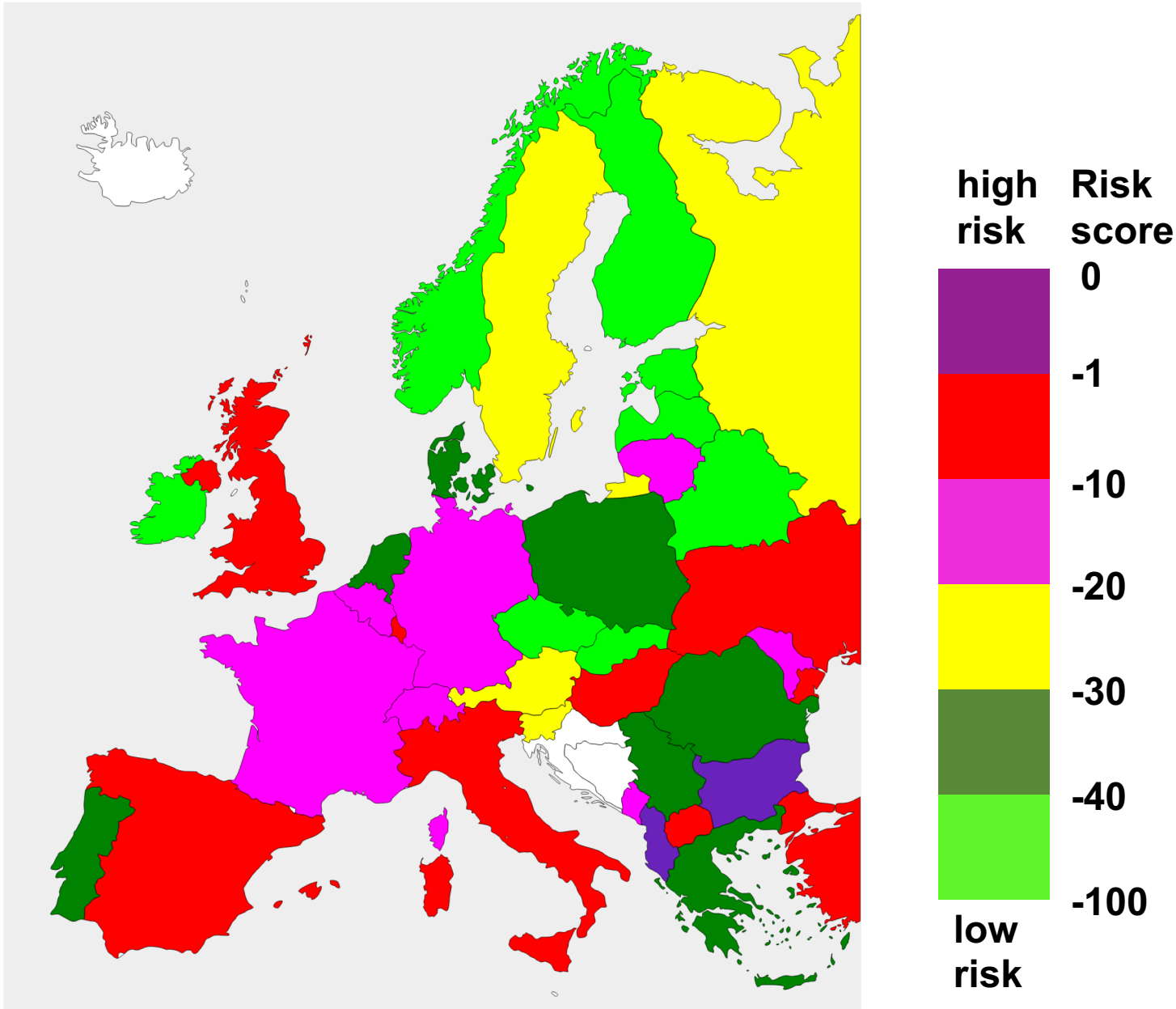


Kamikubo & Takahashi, Fig. 4

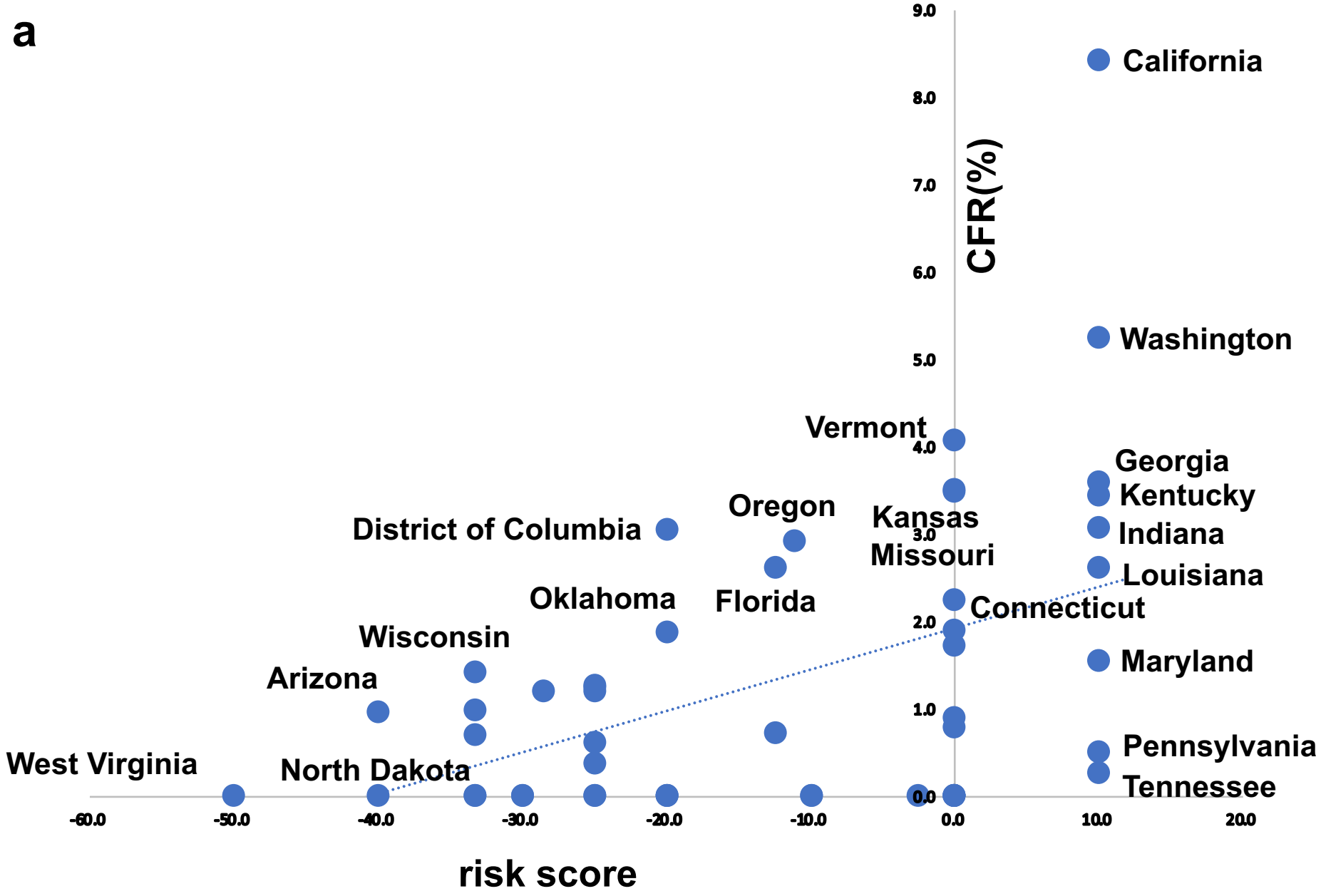


Kamikubo & Takahashi, Fig. 5

b

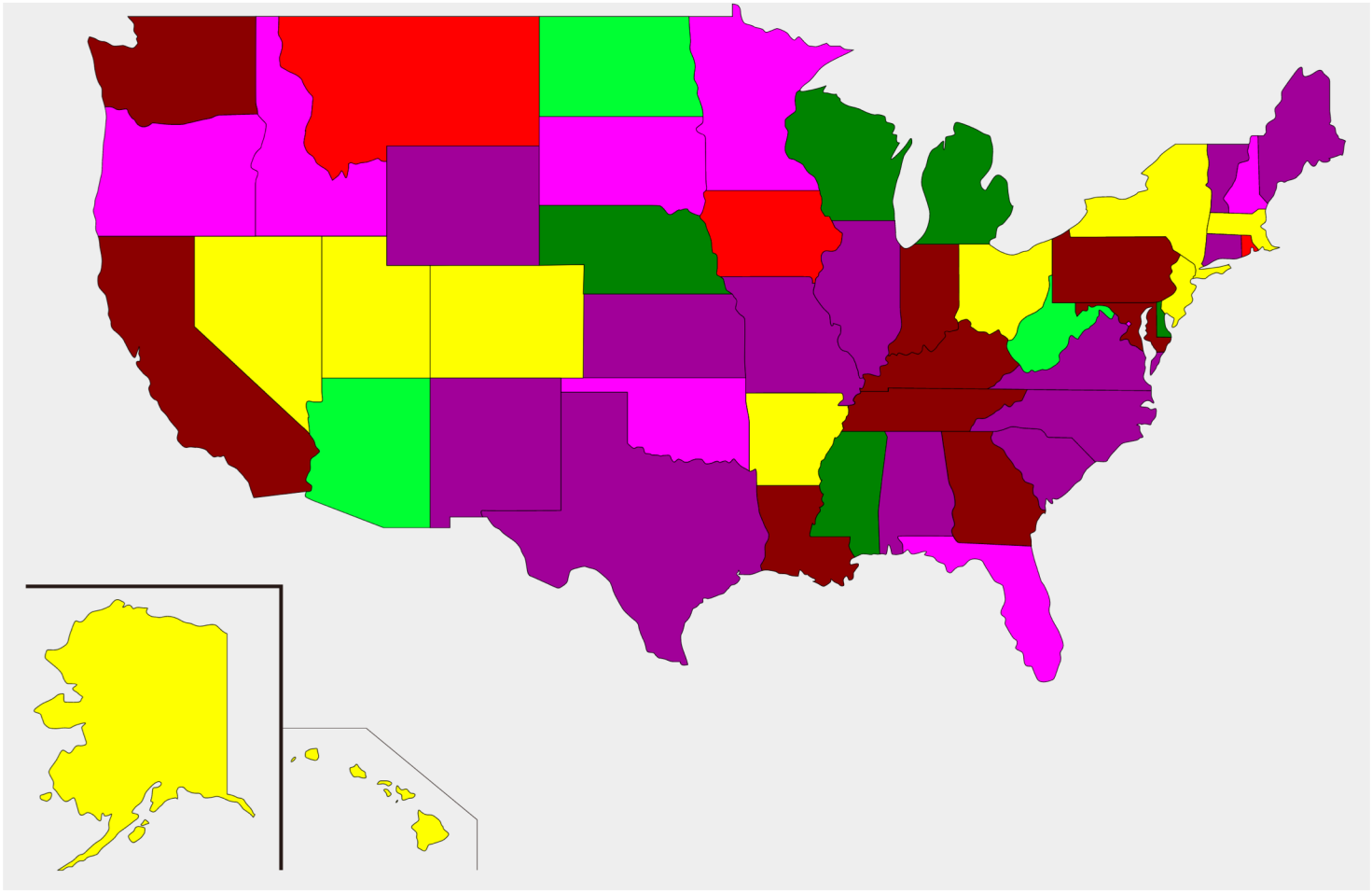


Kamikubo & Takahashi, Fig. 5



Kamikubo & Takahashi, Fig. 6

b



high risk

Risk score

10

0

-10

-20

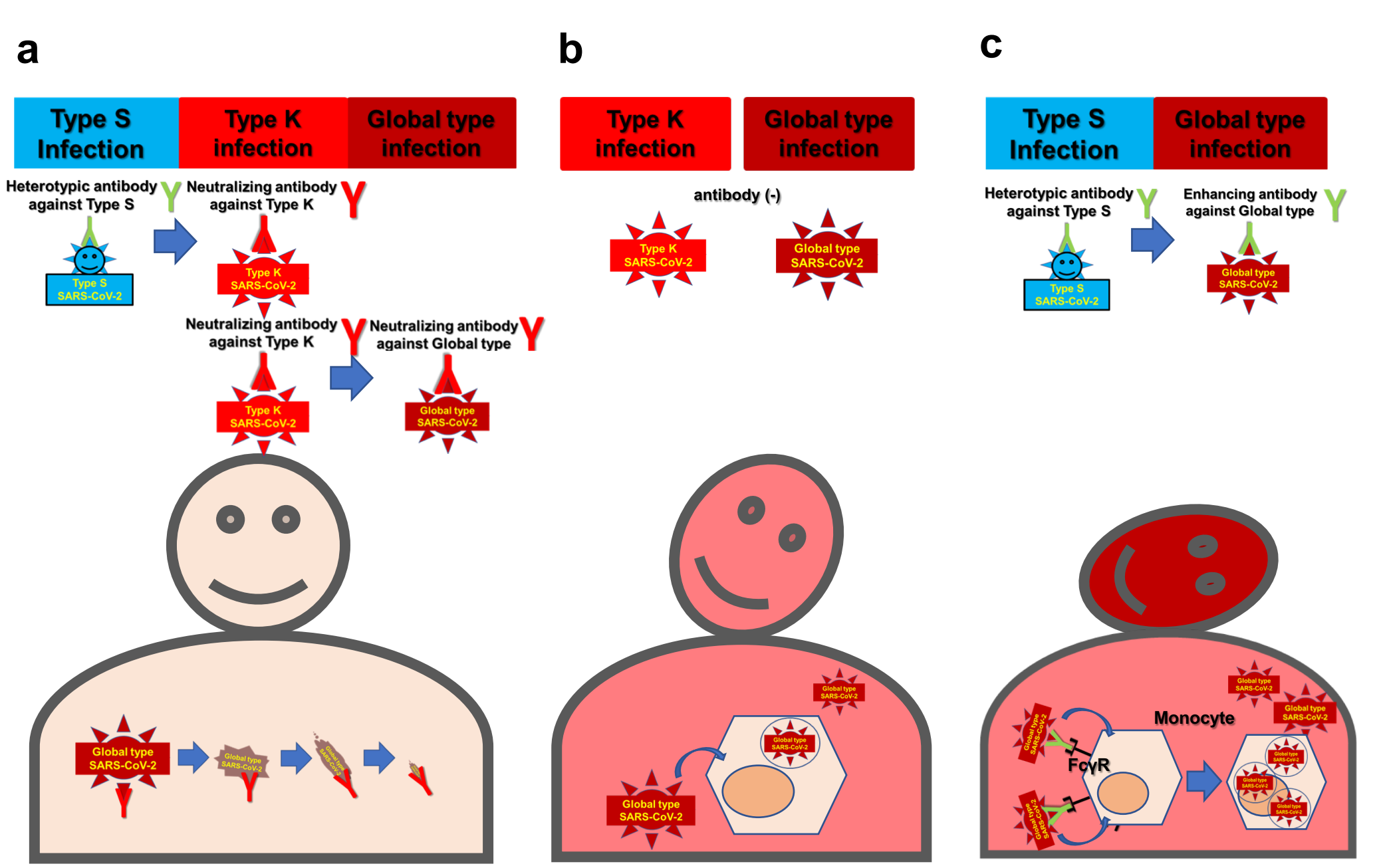
-30

-40

-100

low risk

Kamikubo & Takahashi, Fig. 6



Kamikubo & Takahashi, Fig. 7