"Risk of hospitalization of diagnosed COVID-19 cases during the pandemic: a time-series analsys to unveil short- and long-run dynamics"

Manuela Alcañiz, Marc Estevez and Miguel Santolino

UBIREA | Institut de Recerca en Economia Aplicada Regional i Pública

UBIREA | Institut de Recerca en Economia Aplicada Regional i Pública
UNIVERSITAT DE BARCELONA

WEBSITE: www.ub.edu/irea/ • CONTACT: irea@ub.edu

The Research Institute of Applied Economics (IREA) in Barcelona was founded in 2005, as a research institute in applied economics. Three consolidated research groups make up the institute: AQR, RISK and GiM, and a large number of members are involved in the Institute. IREA focuses on four priority lines of investigation: (i) the quantitative study of regional and urban economic activity and analysis of regional and local economic policies, (ii) study of public economic activity in markets, particularly in the fields of empirical evaluation of privatization, the regulation and competition in the markets of public services using state of industrial economy, (iii) risk analysis in finance and insurance, and (iv) the development of micro and macro econometrics applied for the analysis of economic activity, particularly for quantitative evaluation of public policies.

Abstract

Objectives: With the outbreak of the SARS-CoV-2 pandemic, the unprecedented rise in demand for hospital care brought health systems worldwide to the brink of collapse. The dynamics of the COVID-19 pandemic have alternated periods of high incidence with others of low incidence, making it difficult to separate short- and long-run relationship between the number of COVID-19 cases diagnosed and the demand for hospital beds. The aim of this study is to model the risk of hospitalization of diagnosed cases during the pandemic. Methods: Time series techniques are applied to evaluate the short- and long-run relationship between daily number of COVID-19 cases diagnosed and daily number hospital admissions. Drawing on daily Spanish data from 11 May 2020 to 20 March 2022, we propose an error correction model that introduces a short-run mechanism to adjust transitory disequilibrium in the long term. The impact of the Omicron variant and vaccination on the need for in-patient care are assessed. To examine changes during different life stages, the same analysis is performed by age group. Results: Dynamics between the number of positive cases and demand for hospital beds tends to the equilibrium in the long run, with 9% of any deviation being corrected after one period. Individuals aged between 50 and 69 have benefited most from the reduced severity of the Omicron variant, while vaccination had proved to be less effective for people aged over 80. Conclusions: Models discriminating between the short- and long-run dynamics provide health planners with a valuable demand forecasting tool which should be useful for developing both structural programs and emergency interventions.

JEL classification: C13, C32, I10, I18.

Keywords: COVID-19, Hospitalization risk, Vaccination, Error Correction Model, Health Planning.

Manuela Alcañiz: Riskcenter-IREA, Dept. Econometrics, Statistics and Applied Economy, Universitat de Barcelona (UB). Email: malcaniz@ub.edu

Marc Estévez: Riskcenter-IREA, Dept. Econometrics, Statistics and Applied Economy, Universitat de Barcelona (UB). Email: marcestevez18@gmail.com

Miguel Santolino: Riskcenter-IREA, Dept. Econometrics, Statistics and Applied Economy, Universitat de Barcelona (UB). Email: msantolino@ub.edu

INTRODUCTION

The SARS-CoV-2 pandemic marked a turning point in health planning worldwide. The limited capacity of healthcare and hospital resources and the unforeseen levels of demand resulted in situations bordering on collapse at the most critical moments of the crisis, forcing healthcare services to increase the availability of hospital beds without their being able to implement correct processes of planning [1,2]. In response to these unprecedented circumstances, the scientific literature has hastened to provide health planners with methods for predicting hospital demand at each specific moment of the developing pandemic. Several authors have employed an econometric methodology to model the relationship between the incidence of the disease and the resulting need for hospital in-patient care. Disease incidence and hospital demand recorded during an observation period are both time indexed sequences of data points, so time series techniques are required to analyze their association. Santolino et al. [3], for example, proposed a regression model with multiplicative structure to predict the number of hospitalizations likely to be required nine days in advance. Other authors have studied how the hospital care required by patients with COVID-19 during the pandemic has impacted the number of admissions for other pathologies [4,5].

The aim of this paper is to model the relationship between the number of positive cases of COVID-19 infection and the number of patients hospitalized because of the disease by drawing on Spanish data. The dynamics of the COVID-19 pandemic alternates periods of high incidence with others of low incidence, making it difficult to separate short- and long-run relationship between the number of COVID-19 cases diagnosed and the demand for hospital beds. Given that the number of COVID-19 cases detected and the number of hospital admissions recorded exhibit a common stochastic trend consistent with cointegration, we opt to define an Error Correction Model (ECM) so as to better adjust the long-term equilibrium detected. The ECM separates short- and long-run relationship between detected COVID-19 cases and hospital admissions. We evaluate the potential impact on the long-relationship of the SARS-CoV-2 Omicron variant [6],

as well as that of increasing vaccination rates among the Spanish population. Further, by introducing improvements in the short-run dynamics, we aim to correct the transitory deviations that may present themselves and, in this way, validate the model proposed. Finally, the analysis is carried out by different age groups to determine whether there are any differences of note in the relationship between the incidence of the disease during different life stages and the need for in-patient care.

Particular attention is paid here to estimating the speed at which the relationship between the number of diagnosed cases and the demand for hospital beds returns to its long-run equilibrium after a change in incidence of the coronavirus. This estimation should constitute a valuable contribution to health planning, as it can help shed new light on how best to improve predictions of hospital demand, which has been shown to be subject to stochastic shocks that are not easily anticipated.

Some studies have used ECM to measure the impact of the spread of SARS-CoV-2 on the healthcare system [7-9]. Such models can be useful in healthcare management for anticipating marked increases in the demand for care services and responding appropriately. For example, Nguyen et al. [10] drew on local data from the metropolitan area of Charlotte (United States) to estimate a vector ECM for studying the relationship between the daily infection incidence and the aggregate number of hospital beds occupied by SARS-CoV-2 patients, while Mills [11] explored the changing relationship between infections, hospital admissions and deaths using data from England. However, to date, there have been few attempts to explain the number of hospital admissions in relation to the daily SARS-CoV-2 infection incidence.

METHODS

Time-series data

Two free-access datasets from official organizations are used in this study. The daily number of detected cases and hospital admissions are obtained from Spain's National Epidemiology Centre (https://cnecovid.isciii.es). Positive cases are registered by date of diagnosis and hospitalizations by date of admission. Information is disaggregated by gender, age intervals and province of residence. The percentage of the population fully vaccinated against COVID-19 is obtained from the weekly reporting data of the number of doses administered by age groups provided by the European Centre for Disease Prevention and Control (https://opendata.ecdc.europa.eu/). Our series covers the time period from 11 May 2020 to 20 March 2022. Records of positive cases and hospital admissions are actually available from 1 January 2020, but these initial months are excluded because of the low diagnostic capacity attributable to severe testing restrictions [12]. Moreover, the criteria for registering positive cases were modified as of the 10 May 2020 [3].

Our preliminary analysis of the series revealed that the number of positive cases and hospital admissions presented a multiplicative weekly seasonality with cases being underreported at weekends. A log transformation was applied to both time series and the seasonal effect was adjusted. Seasonal and trend decomposition using the Loess method (STL) was applied in order to decompose the time series [13]. Weekly vaccination information was converted to a daily time series assuming that the same number of doses was administered daily throughout the week. Figure 1 plots the positive cases, hospital admissions and the percentage of population fully vaccinated against COVID-19. For comparison purposes, positive cases and hospital admissions are shown on a 0-100 scale in Figure 1.

[INSERT FIGURE 1]

Fig. 1 Time series* for COVID-19 detected positives, hospital admissions and percentage of population fully vaccinated in Spain for the period from 11 May 2020 to 20 March 2022

Note: * For comparison purposes, positive cases and hospital admissions were transformed on a standardized scale from 0 to 100.

Short- and long-run relationships

When two time series present a long-run relationship their dynamic common behaviour can be described using an error correction model [14]. Such a model includes a term that accounts for deviations from that long-run relationship and provides an estimate of the speed at which the disequilibrium dissipates. When multiple time series are analysed jointly their order of integration needs to be examined, which is usually achieved using the augmented Dickey-Fuller (ADF) test [15]. If non-stationary variables are included in a regression model, spurious outcomes may occur. When these non-stationary variables are integrated of order 1 and there is a linear relationship, such that their residuals are stationary (cointegration), an ECM may constitute an appropriate methodological approach [16].

Given the cointegration here between our two time series – that is, positive cases of COVID-19 and hospital admissions – we propose employing an ECM. This model can link the long-run equilibrium between the two time series with the short-run adjustment mechanism that describes how the relationship reacts to stochastic fluctuations in the incidence of the disease.

The long-run equilibrium relationship between positive cases and number of hospitalizations is represented by the cointegration equation as follows:

$$y_t = b_0 + b_1 x_t + b_2 x_t I_{omic,t} + b_3 z_t + ect_t \quad (1)$$

where y_t corresponds to the logarithm of new hospital admissions on day *t* and x_t is the logarithm of the number of daily positive COVID-19 cases and $t=1,\ldots,T$, where T=679, given that this is the number of days in the period under study. To analyse the effect of the Omicron variant on the

long-term relationship between the number of hospitalizations and positive cases, the dummy variable $I_{omic,t}$ takes a value of 1 if *t* occurs on or after 29 November 2021 ($t \geq 112$), the earliest date from which the Omicron variant was detected in Spain, and zero otherwise. The constant term is b_0 and z_t indicates the percentage of fully vaccinated population at time *t*. Finally, the error correction term (ect) captures the regression residuals.

If the residuals in (1) are stationary, the variables are cointegrated. An ECM can then be specified to analyse the short-run adjustment mechanism and the long-run equilibrium between these variables as follows:

$$\Delta y_t = c + \sum_{i=1}^{k} \psi_i \Delta y_{t-i} + \sum_{j=0}^{q} w_j \Delta x_{t-j} + \gamma \cdot ect_{t-1} + \varepsilon_t \quad (2)$$

In our case, the first difference of log hospital admissions ($\Delta y_t = y_t - y_{t-1}$) is regressed on the lagged error correction term from (1), *k*-lagged values of the same variable and current and *q*-lagged values of the log hospital admissions, all in differences. The optimal numbers for *k* and *q* are determined when estimating the model. Coefficients ψ_i and w_j measure short-run reactions of the dependent variable with its previous changes and with changes in the explanatory variables, respectively. The intercept included in the regression is *c*, while γ corresponds to the error correction rate that indicates the speed of adjustment in the short term when there is a disequilibrium in the long term, i.e., $ect_{t-1} \neq 0$ [17]. Finally, ε_t is the error term which is normally distributed with zero mean and variance σ_t^2, $\varepsilon_t \sim N(0, \sigma_t^2)$. A generalized autoregressive conditional heteroscedasticity (GARCH) [18] model specification is used to deal with the presence of heteroscedasticity. A GARCH(1,1) is proposed here to model the variance as $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2$. To conclude, it should be borne in mind that an autoregressive distributed lag model specification could be obtained from (2) by rearranging the variables [19].

RESULTS

Error correction model

The statistical analysis was conducted using R statistical software, version 4.1.1. [20,21]. The first step in this analysis involved examining the order of integration of the series. The values of the ADF test statistics for the logarithm of hospital admissions y_t and the logarithm of positive cases x_t were -0.031 and 0.453, respectively. As a result, the null hypothesis, which states the presence of a unit root, was not rejected at a significance level of 5%. However, the null hypothesis was rejected when Δx_t and Δy_t were analyzed, indicating that the first difference of the time series were stationary (ADF(Δy_t)=-23.094 and ADF(Δx_t)=-23.660). Threfore, both variables x_t and y_t are integrated of order one.

Log-run equilibrium

The cointegration equation expressed in (1) is estimated using fully modified least squares [22]. Table 1 reports the coefficient estimates and the ADF test performed on the cointegration residuals. The results show that the residuals are integrated of order zero (stationary), thus cointegration exists.

Table 1. Estimation of the cointegrating equation (long-run relationship) between time series of positive COVID-19 cases and hospital admissions, and ADF test on residuals

Coefficient	Description	Estimate
b_0	Intercept	-1.228**
b_1	Positive cases (log)	0.855**
b_2	Omicron variant	-0.076**
b_3	% vaccinated population	-0.008**
ADF	ADF test on *ect*	-5.413**

Note: ** p-value < 0.01.

The long-run coefficient of the (log) number of positive cases is greater than 0, meaning that an increase in the number of COVID-19 cases diagnosed implies an increase in the number of hospitalizations. However, the long-run coefficient for the percentage of fully vaccinated is negative, which indicates that the greater the number of people with full vaccination status, the fewer the number of people that have to be hospitalized. Additionally, the coefficient associated with the Omicron variant presents a significant negative sign, suggesting that after 29 November 2021 an increase in the number of positive cases is associated with a smaller increase in the number of patients requiring hospitalization. However, if first we only consider the period before the appearance of the Omicron variant ($I_{omic,t} = 0$), the long-run relationship between hospital admissions and positive cases can be represented as follows:

$$e^{E[y_t]} = e^{-1.228+0.855x_t-0.008z_t} = \frac{e^{-1.228}(e^{x_t})^{0.855}}{(e^{0.008})^{z_t}}, \text{ i.e., } Hos_t = 0.293\frac{{Pos_t}^{0.855}}{1.008^{z_t}},$$

where Hos_t indicates the exponential of the expected log number of hospitalizations in *t*, $e^{E[y_t]}$, and Pos_t indicates the number of positive cases, $Pos_t = e^{x_t}$. This means that to estimate the number of hospital admissions, the number of positive cases must be raised to 0.855 and then divided by 1.008^{z_t}. Thus, 29.3% of this corrected number of positive cases $\left({Pos_t}^{0.855}/1.008^{z_t}\right)$ is estimated as being admitted to hospital. After 29 November 2021 ($I_{omic,t} = 1$), the log-run relationship is:

$$Hos_t = e^{-1.228+(0.855-0.076)x_t-0.008z_t} = 0.293\frac{{Pos_t}^{0.779}}{1.008^{z_t}}.$$

Error correction model estimation

The estimated coefficients of the ECM corrected for heteroscedasticity through a GARCH(1,1) specification are shown in Table 2. The selection of the order (*k*,*q*) was based on the Bayesian information criterion (BIC) [23]. The model specification with the lowest BIC had 11 lags on the difference of log hospital admissions and 1 lag on the difference of log positive cases.

Table 2. Error correction model (short-run relationship) between time series of positive COVID-19 cases and hospital admissions (in log scale)

Coefficient	Description	Estimate
c	Intercept	-0.003
ψ_1	1-lagged hospitalization difference	-0.591**
ψ_2	2-lagged hospitalization difference	-0.298**
ψ_3	3-lagged hospitalization difference	-0.074
ψ_4	4-lagged hospitalization difference	0.009
ψ_5	5-lagged hospitalization difference	0.115*
ψ_6	6-lagged hospitalization difference	0.160**
ψ_7	7-lagged hospitalization difference	0.349**
ψ_8	8-lagged hospitalization difference	0.232**
ψ_9	9-lagged hospitalization difference	0.183**
ψ_{10}	10-lagged hospitalization difference	0.179**
ψ_{11}	11-lagged hospitalizations difference	0.105**
w_0	Difference of positive cases	0.304**
w_1	1-lagged difference of positive cases	-0.047**
Error correction		
γ	Error correction term	-0.088**
Variance equation		
α_0	Variance equation intercept	$1.6 \cdot 10^{-4}$**
α_1	Variance equation error term	0.110**
α_2	Variance equation variance term	0.866**
AIC	AIC of the ECM	-2.169
BIC	BIC of the ECM	-2.049
HC	HQ of the ECM	-2.123
R^2	ECM coefficient of determination	0.425

Note: ** p-value < 0.01; * p-value < 0.05.

Table 2 shows that the coefficient associated with the error correction term is significant and takes a value between -1 and 0, which are the necessary conditions for stating that the dynamics between the analysed variables tend to equilibrium. Specifically, the estimated coefficient reflects the speed of adjustment in case of long-run disequilibrium. Its value suggests that around 9% of any

deviation from the long-run relationship dissipates after one period. In the case of the short-run coefficients, the 1- and 2-period lagged differences of the (log) number of hospitalizations have a negative impact on the expected difference of the (log) number of hospital admissions. However, their impact is positive when the 5- to 11-period lagged differences of the (log) number of hospitalizations are considered. Finally, the difference between the (log) number of positive cases and this series lagged one period serves to explain the difference in the (log) number of hospitalizations.

Model diagnostics

To obtain both consistent and efficient estimates, the residuals in (2) should follow an uncorrelated white noise process. Figure 2 shows the partial autocorrelations of the model (2) residuals. The rejection limits of the null hypothesis stating that the residuals follow a white noise process are computed under the independent and identically distributed (IID) and GARCH assumptions [24]. Under the IID hypothesis, the partial autocorrelations for 1, 7, 17 and 24 lags are significant. However, these partial autocorrelations are not-significant when a GARCH process is considered, which leads to the conclusion that the null hypothesis cannot be rejected.

[INSERT FIGURE 2]

Fig. 2 Partial autocorrelation function of residuals and rejection limits of white noise process under IID and GARCH hypotheses

The adjusted Pearson goodness-of-fit test can be used to check if the selected conditional error distribution is appropriate [25]. This test is based on the chi-squared goodness-of-fit test which compares the empirical distribution of standardized and theoretical residuals. For all the bins considered when classifying the values (from 20 to 50, from ten at a time), the normal conditional distribution of residuals in the ECM with GARCH(1,1) was not rejected at the 5% significance level.

The Ljung-Box test of standardized residuals analyses the serial dependence for the residuals of the mean process [25]. The null hypothesis of no serial correlation was not rejected with either 1 lag (p-value: 0.345), 32 lags (p-value: 0.999) or 54 lags (p-value: 0.723), suggesting that the ECM specification for the conditional mean process is adequate. The Ljung-Box test on *squared* standardized residuals and the auto-regressive conditional heteroscedasticity (ARCH)-LM test provide a means of detecting a time-varying phenomenon in the residuals of the variance process. The null hypothesis of no autocorrelation was not rejected at the 5% of significance level with either 1, 5 or 9 lags in the Ljung-Box test, nor with 1, 5 and 7 lags in the ARCH-LM test. Thus, it can be concluded that the ECM with GARCH(1,1) captures the dynamics of the variance process.

The sign bias test proposed by Engle and Ng [26] is useful for evaluating the presence of leverage effects by regressing the squared standardized residuals on lagged negative and positive shocks. The null hypothesis of no leverage effects was not rejected at the 5% significance level; hence, there was no evidence of misspecification of the conditional variance process.

Age groups

The ECM was further calibrated for different age groups with a twofold goal: 1) to evaluate whether there are differences in the way the number of positive cases and vaccination status affect the number of hospital admissions and, 2) to compare the speed of adjustment in the case of long-run disequilibrium. The age intervals considered are 20–49, 50–69, 70–79 and 80 years or more[1]. The results of the estimation of the ECM and the cointegration equation for each age group are shown in Table

3. The selection of the number of lags in each ECM was based on BIC. It is evident that the number of lags falls drastically compared to results of the ECM for the whole population. The

[1] Information on the number of positive cases and the number of hospitalizations was specifically available for these age groups. In the case of the fully vaccinated population, information was available for the following age groups: 25–49, 50–69, 70–79 and 80 years or more.

selected models now include two/three lags of the difference of (log) hospitalizations and zero/one lag of the difference of (log) positive cases.

Table 3. Error correction model for positive cases and hospital admissions (in log scale), and long-run coefficients by age groups

Coeff.	Description	Estimates			
		20-49	50-69	70-79	80+
Long-run coefficients					
b_0	Intercept	-1.970**	-2.198**	-1.372**	-1.004**
b_1	Positive cases (log)	0.829**	1.005**	1.005**	0.998**
b_2	Omicron variant	-0.114**	-0.146**	-0.133**	-0.117**
b_3	% vaccinated population	-0.005**	-0.005**	-0.004**	-0.001*
Short-run coefficients					
c	Intercept	-0.001	-	-	0.001
ψ_1	1-lagged hospitalization difference	-0.665**	-0.637**	-0.776**	-0.548**
ψ_2	2-lagged hospitalization difference	-0.316**	-0.266**	-0.433**	-0.234**
ψ_3	3-lagged hospitalization difference	-	-	-0.162**	-
w_0	Difference of positive cases	0.301**	0.365**	0.410**	0.314**
w_1	1-lagged difference of positive cases	-	-	-	0.091**
Error correction					
γ	Error correction term	-0.076**	-0.072**	-0.112**	-0.112**
Variance equation					
α_0	Variance equation intercept	$4\cdot10^{-4*}$	$5\cdot10^{-4**}$	$4\cdot10^{-4*}$	$2\cdot10^{-4*}$
α_1	Variance equation error term	0.126**	0.110**	0.126**	0.121**
α_2	Variance equation variance term	0.861**	0.866**	0.868**	0.878**
AIC	AIC of the ECM	-0.858	-1.126	-0.717	-1.008
BIC	BIC of the ECM	-0.804	-1.072	-0.656	-0.948
HQ	HQ of the ECM	-0.837	-1.105	-0.693	-0.985
R^2	ECM coefficient of determination	0.436	0.409	0.478	0.290

Note: ** significance level at 1%; * significance level at 5%.

A number of differences by age group in the long-run relationship and the speed of adjustment in the short-run are worth highlighting. For example, in the case of the long-run equilibrium coefficients, the intercept estimates present higher values for the older age groups. Thus, the value e^{b_0}, corresponding to the proportion of corrected positive cases $\left(Pos_t{}^{b_1+b_2 I_{omic,t}}/e^{-b_3 z_t}\right)$ estimated as being admitted to hospital, is higher for older ages. Additionally, the coefficients b_1, associated with the correction of positive cases, indicate that an increase in the number of cases is associated with an increase in the number of hospitalizations; however, this increment is smaller in the case of the youngest age group. This impact on the number of hospitalizations falls after 29 November 2021, as the coefficients b_2, associated with the Omicron variant, are significant and negative in all ages groups, most notably in the 50–69 age group. It is also evident that the beneficial effects of an increase in the percentage of fully vaccinated b_3 on falling numbers of hospitalizations holds for all ages, albeit that this effect decreases as people get older.

Additionally, all the error correction term coefficients are significant and negative. The number of days required to close the gap between current and equilibrium hospital admissions for the different group ages are shown in Figure 3. The number of days required to correct a long-run disequilibrium are computed using the approach devised by Galeotti et al. [27]. These results are obtained using the equation

$$Days = \frac{\ln\left(\frac{y_t - y_t^*}{y_0 - y_t^*}\right)}{\gamma} = \frac{\ln(1 - gap)}{\gamma}$$

where *gap* represents the difference between current and equilibrium (log) hospitalizations defined as $gap = \frac{y_0 - y_t}{y_0 - y_t^*}$, y_t is the logarithm of current hospitalizations, y_0 is the logarithm of initial hospitalizations and y_t^* represents equilibrium (log) hospitalizations. Finally, γ is the error correction term that shows the speed of adjustment to the long-run equilibrium. We find that adjustment to the long-run equilibrium is more rapid in the case of the two oldest age groups (i.e. 70–79 and 80+) than in that of the two youngest (20–49 and 50–69), placing the speed adjustment

line for the general model approximately in the middle. For instance, an 80% gap between current hospitalizations and the number of hospital admissions in the long-run equilibrium takes over 14 days to be closed in the case of the two oldest groups, while approximately 21 days are required for the 20–49 age group, 22 days for the 50–69 age group and 18 days for all ages.

[INSERT FIGURE 3]

Fig. 3 Speed of adjustment to equilibrium in days

These particular results obtained after disaggregating by age should, however, be treated with some caution, given that residual autocorrelation that might affect the efficiency of estimates was detected [25]. This effect in all likelihood is attributable to the low number of lagged regressors selected based on the best goodness-of-fit performance of the models.

DISCUSSION

This article has examined the relationship between the number of COVID-19 cases detected in Spain and the number of hospital admissions due to the virus between May 2020 and March 2022. Short- and long-run dynamics were separated by means of an ECM corrected for heteroscedasticity using a GARCH variance process. The long-run relationship followed a multiplicative model (additive after logarithmic transformations), in line with Santolino et al. [3] who linked hospital admissions and nine-period-lagged positive cases. The dynamics of the relationship between our time series proved to be stable and tended to long-run equilibrium. When a disequilibrium in the long-run relationship was observed, around 9% of the gap was corrected after one period.

The behaviour of the pandemic indicators was not steady over time as periods of high incidence alternated with others of low incidence. We detected two distinct factors impacting the long-run equilibrium between the number of cases and the number of hospitalizations: that is, vaccination and the presence of the Omicron variant. The vaccination program against COVID-19 was initiated in Spain on 27 December 2020. Previous studies have shown that vaccination reduces the risk of hospitalization [28,29]. Our research also shows that hospital pressure decreased as the vaccination program was progressively rolled out. Hence, while a 1% increase in the fully vaccinated population reduced the risk of hospitalization by 0.8%. The Omicron variant appeared in Europe at the end of November 2021, and by January 2022 it was already considered dominant [30]. Some studies suggest that this variant presents a lower risk of hospitalization and death than earlier variants of the virus [31]. Here, we found that the number of positive cases has to be raised to a power to compute the number of hospitalizations. This power value ranged from 0.885 before 29 November 2021 to 0.779 after this date. The fact that both values are lower than one means that any increase in the number of positive cases generates a lower increase in the number of hospitalizations, with this reduction being greater following the appearance of the Omicron variant.

Alternative designs of the long-run equilibrium equation were investigated to capture other forms of relationship between the number of hospitalizations and the number of positives detected, vaccination status and the Omicron variant, but a poorer performance was observed in terms of goodness-of-fit in all cases. The effect of the Delta variant on the number of hospitalizations was also analysed and its effect was found not to be statistically significant at the 5% level in any model design. However, we did not find any effect of the Omicron and Delta variants on the short-run dynamics of the number of hospitalizations. In fact, the difference in number of hospitalizations was essentially explained by the previous (lagged) differences of the hospitalization time series.

COVID-19 hospitalization rates are known to be exponentially associated with age [32]. In this study, we analysed the long- and short-run relationships between positive cases and hospital admissions for four different age groups. Results show that the percentage of individuals with a positive diagnosis requiring hospitalization was higher for those aged over 50. The appearance of the Omicron variant seemed to reduce the severity of the disease in all groups, but particularly among those aged 50 to 69. Sievers et al. [33], likewise, worked with different age groups and found this reduction to be apparently greater for those aged over 35. Finally, in line with other studies [34,35], our results showed that vaccination proved to be less effective for older people. As for the speed of correction in the presence of deviations, we observed that the return to the long-run equilibrium path by way of short-run adjustments was faster for the two oldest age groups.

To date, the different waves of the virus have produced fluctuating and unpredictable levels of pressure on hospitals [36,37]. At given moments the resources available have been insufficient, bringing hospitals to the brink of collapse, with an obvious detrimental impact for patients and healthcare personnel alike. The dynamic model proposed here seeks to be a useful health planning tool that can forecast the amount of hospital resources required at any specific moment based on the prevailing incidence of the disease, the virulence of the dominant variant, and the proportion of the population with full vaccination status. By differentiating between the long- and short-run effects, our methodology furnishes healthcare decision-makers with a dual mechanism that facilitates their evaluations of the impact of (i) the structural health policies aimed at addressing the long-run relationship between positive cases of COVID-19 and the demand for hospital beds, and (ii) emergency interventions with a short-run impact on the demand for hospital admissions.

This study is not exempt from limitations, not least the fact that access to reliable data is essential for developing an accurate, realistic model. As Hyafil and Moriña [38] stress, the number of tests performed has a direct effect on the number of positive cases detected, which suggests there are likely to have been undiagnosed cases not considered in this study. Additionally, here we have

had to use weekly vaccination information to estimate daily rates of inoculation, as complete daily data on the number of individuals with full vaccination status were unavailable. Finally, risk factors, other than age, were not analysed as they were not registered by the databases. For example, the effect of gender on hospitalization risk was not investigated, although some studies suggest that being male represents a risk factor for COVID-19 [39]. All in all, this study highlights the importance of having reliable, homogeneous, disaggregated and up-to-date information when evaluating the behaviour of indicators of such great interest for public health.

CONCLUSIONS

The application of the error correction model introduces a short-run mechanism to adjust transitory disequilibrium in the long-run relationship between the number of COVID-19 cases diagnosed and the consequent demand for hospital beds. Results reveal that the dynamics between the number of positive cases and demand for hospital beds tends to the equilibrium in the long run, with 9% of any deviation being corrected after one period. The dynamic modelling approach proposed herein should represent a valuable instrument for planning hospital resources in any pathology that necessitates in-patient care, especially epidemics with waves of contagion. Based on the cases detected in primary care, it would be possible to predict the number of hospital admissions, thus allowing health planners to anticipate both the long- and short-run impact on hospital pressure created by any disease.

References

1. Peiffer-Smadja N, Lucet J-C, Bendjelloul G, Bouadma L, Gerard S, Choquet C, et al. Challenges and issues about organizing a hospital to respond to the COVID-19 outbreak: experience from a French reference centre. Clin Microbiol Infect. 2020;26(6):669-72. https://doi.org/10.1016/j.cmi.2020.04.002.

2. Lee A, English P, Pankhania B, Morling JR. Where England's pandemic response to COVID-19 went wrong. Public health. 2021;192:45–8. https://doi.org/10.1016/j.puhe.2020.11.015.

3. Santolino M, Alcañiz M, Bolancé C. Hospitalization from covid-19: a health planning tool. Rev Saude Publica. 2022;56:51. https://doi.org/10.11606/s1518-8787.2022056004315.

4. Birkmeyer JD, Barnato A, Birkmeyer N, Bessler R, Skinner J. The Impact Of The COVID-19 Pandemic On Hospital Admissions In The United States. Health Aff. 2020;39(11):2010-7. https://doi.org/10.1377/hlthaff.2020.00980.

5. McIntosh A, Bachmann M, Siedner MJ, Gareta D, Seeley J, Herbst K. Effect of COVID-19 lockdown on hospital admissions and mortality in rural KwaZulu-Natal, South Africa: interrupted time series analysis. BMJ Open. 2021;11(3):e047961. https://doi.org/10.1136/bmjopen-2020-047961.

6. Tian D, Sun Y, Xu H, Ye Q. The emergence and epidemic characteristics of the highly mutated SARS-CoV-2 Omicron variant. J Med Virol. 2022;94(6):2376-83. https://doi.org/10.1002/jmv.27643.

7. Berta P, Paruolo P, Verzillo S, Lovaglio PG. A bivariate prediction approach for adapting the health care system response to the spread of COVID-19. PLoS ONE. 2020;15(10):e0240150. https://doi.org/10.1371/journal.pone.0240150.

8. Monllor P, Su Z, Gabrielli L, Taltavull de La Paz P. COVID-19 Infection Process in Italy and Spain: Are Data Talking? Evidence From ARMA and Vector Autoregression Models. Front Public Health. 2020;8:550602. https://doi.org/10.3389/fpubh.2020.550602.

9. González-Pérez B, Núñez C, Sánchez JL, Valverde G, Velasco JM. Expert System to Model and Forecast Time Series of Epidemiological Counts with Applications to COVID-19. Mathematics. 2021;9(13):1485. https://doi.org/10.3390/math9131485.

10. Nguyen H, Turk P, McWilliams A. Forecasting COVID-19 Hospital Census: A Multivariate Time-Series Model Based on Local Infection Incidence. JMIR Public Health Surveill. 2021;7(8):e28195. https://doi.org/10.2196/28195.

11. Mills TC. Modelling the link between Covid-19 cases, hospital admissions and deaths in England. National Accounting Review. 2022;4(1):38-55. https://doi.org/10.3934/NAR.2022003.

12. Sidawi T, Garau J. Mortalidad de COVID-19. Tendencias en la evolución de la pandemia. Med Balear. 2021;36:42-9. https://doi.org/10.3306/MEDICINABALEAR.36.01.42.

13. Cleveland R, Cleveland W, McRae J, Terpenning I. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. J Off Stat. 1990;6:3-73.

14. Engle RF, Granger CWJ. Co-integration and Error Correction: Representation, Estimation, and Testing. Econometrica. 1987;55:251-76.

15. Said SE, Dickey DA. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika. 1984;71(3):599-607. https://doi.org/10.2307/2336570.

16. Asteriou D, Hall SG. Applied Econometrics. 3rd ed. London: Palgrave Macmillan; 2015.

17. Alogoskoufis G, Smith R. On error correction models: specification, interpretation, estimation. J Econ Surv. 1991;5(1):97-128. https://doi.org/10.1111/j.1467-6419.1991.tb00128.x.

18. Bollerslev T. Generalized autoregressive conditional heteroscedasticity. J Econom. 1986;31(3):307-27. https://doi.org/10.1016/j.eswa.2022.117951.

19. Hassler U, Wolters J. Autoregressive distributed lag models and cointegration. Allg stat Arch. 2006;90:59–74. https://doi.org/10.1007/s10182-006-0221-5.

20. R Core Team. R: A language and environment for statistical computing. Version 4.1.1 (software). R Foundation for Statistical Computing, Vienna, Austria. 2021. Available from: https://www.R-project.org/.

21. Ghalanos A. rugarch: Univariate GARCH models. R package version 1.4-8. 2022. Available from: http://cran.r-project.org/web/packages.

22. Phillips PCB, Hansen B. Statistical Inference in Instrumental Variables Regression with I(1) Processes. Rev. Econ. Stud. 1990;57(1):99-125, https://doi.org/10.2307/2297545.

23. Schwarz G. Estimating the dimension of a model. Ann Stat. 1978;6(2):461–4. https://doi.org/10.1214/aos/1176344136.

24. Francq C, Zakoian JM. GARCH models: structure, statistical inference and financial applications. 1st ed. New Jersey: John Wiley & Sons; 2010.

25. Palm FC. GARCH Models of Volatility. Handb Stat. 1996;14:209-35.

26. Engle RF, Ng VK. Measuring and Testing the Impact of News on Volatility. J Finance. 1993;48(5):1749-78. https://doi.org/10.1111/j.1540-6261.1993.tb05127.x.

27. Galeotti M, Lanza A, Manera M. Rockets and Feathers Revisited: An International Comparison on European Gasoline Markets. Energy Economics. 2003;25(2):175–90. https://doi.org/10.1016/S0140-9883(02)00102-0.

28. Moline HL, Whitaker M, Deng L, Rhodes JC, Milucky J, Pham H, et al. Effectiveness of COVID-19 vaccines in preventing hospitalisation among adults aged ≥65 years — COVID-NET, 13 States, February–April 2021. MMWR Morb Mortal Wkly Rep. 2021;70:1088–93. https://doi.org/10.15585/mmwr.mm7032e3.

29. Bajema KL, Dahl RM, Prill MM, Meites E, Rodriguez-Barradas MC, Marconi VC, et al. Effectiveness of COVID-19 mRNA Vaccines Against COVID-19–Associated Hospitalization —

Five Veterans Affairs Medical Centers, United States, February 1–August 6, 2021. MMWR Morb Mortal Wkly Rep 2021;70:1294–9. http://dx.doi.org/10.15585/mmwr.mm7037e3.

30. European Centre for Disease Prevention and Control. Assessment of the further spread and potential impact of the SARS-CoV-2 Omicron variant of concern in the EU/EEA, 19th update. ECDC: Stockholm; 2022. https://www.ecdc.europa.eu/en/publications-data/covid-19-Omicron-risk-assessment-further-emergence-and-potential-impact. Accessed Jun 2022.

31. Ulloa AC, Buchan SA, Daneman N, Brown KA. Estimates of SARS-CoV-2 Omicron Variant Severity in Ontario, Canada. JAMA. 2022;327(13):1286–8. https://doi.org/10.1001/jama.2022.2274.

32. Palmer S, Cunniffe N, Donnelly R. COVID-19 hospitalization rates rise exponentially with age, inversely proportional to thymic T-cell production. J R Soc Interface. 2021;18(176):20200982. https://doi.org/10.1098/rsif.2020.0982.

33. Sievers C, Zacher B, Ullrich A, Huska M, Fuchs S, Buda S, et al. SARS-CoV-2 Omicron variants BA.1 and BA.2 both show similarly reduced disease severity of COVID-19 compared to Delta, Germany, 2021 to 2022. Euro Surveill. 2022;27(22). https://doi.org/10.2807/1560-7917.ES.2022.27.22.2200396.

34. Müller L, Andrée M, Moskorz W, Drexler I, Walotka L, Grothmann R, et al. Age-dependent Immune Response to the Biontech/Pfizer BNT162b2 Coronavirus Disease 2019 Vaccination. Clin Infect Dis. 2021; 73(11):2065-72. https://doi.org/10.1093/cid/ciab381.

35. Yelin I, Katz R, Herzel E, Berman-Zilberstein T, Ben-Tov A, Kuint J, et al. Associations of the BNT162b2 COVID-19 vaccine effectiveness with patient age and comorbidities. medRxiv. 2021. https://doi.org/10.1101/2021.03.16.21253686.

36. McCabe R, Schmit N, Christen P, D'Aeth JC, Løchen A, Rizmie D, et al. Adapting hospital capacity to meet changing demands during the COVID-19 pandemic. BMC Med. 2020;18(1):329. https://doi.org/10.1186/s12916-020-01781-w.

37. Dar M, Swamy L, Gavin D, Theodore A. Mechanical-Ventilation Supply and Options for the COVID-19 Pandemic. Leveraging All Available Resources for a Limited Resource in a Crisis. Ann Am Thorac Soc. 2021;18(3):408-16. https://doi.org/10.1513/AnnalsATS.202004-317CME.

38. Hyafil A, Moriña D. Analysis of the impact of lockdown on the reproduction number of the SARS-Cov-2 in Spain. Gac Sanit. 2021;35(5):453-8. https://doi.org/10.1016/j.gaceta.2020.05.003.

39. Okpechi SC, Fong JT, Gill SS, Harman JC, Nguyen TH, Chukwurah QC, et al. Global Sex Disparity of COVID-19: A Descriptive Review of Sex Hormones and Consideration for the Potential Therapeutic Use of Hormone Replacement Therapy in Older Adults. Aging Dis. 2021;12(2):671-83. https://doi.org/10.14336/AD.2020.1211.

UBIREA
Institut de Recerca en Economia Aplicada Regional i Pública
Research Institute of Applied Economics
Universitat de Barcelona
Av. Diagonal, 690 • 08034 Barcelona

WEBSITE: www.ub.edu/irea • **CONTACT:** irea@ub.edu