BBVA Research

Creating Opportunities

N.º 22/03

Working Paper

Covid-19 in Spain during 2021: what have vaccines achieved and what is the health cost of vaccine hesitancy?

José E. Boscá, José Cano, Javier Ferri

May 2022

Covid-19 in Spain during 2021: what have vaccines achieved and what is the health cost of vaccine hesitancy?*

José E. Boscá†

University of Valencia and Fedea

José Cano‡

University of Valencia

Javier Ferri§

University of Valencia and Fedea

May, 2022

Abstract

We evaluate the evolution of the epidemic in Spain during 2021. First, we report significant differences in apparently similar health series depending on the data source. Second, we measure the trend in pressure on hospitals and intensive care units (ICUs). Third, we provide a picture of infections corrected for their capacity to generate severe disease and hospital pressure. Fourth, we estimate the pressure of the latest surge on primary medical care, in terms of the increase in health care staff needed to keep the workload per employee constant at pre-pandemic levels. Fifth, we estimate the effect of vaccination in terms of reducing infections, hospitalizations, ICU admissions, and deaths. In relation to the effect of vaccines, we evaluate the health cost of anti-vaxxers during the sixth wave. We also compare the health impact of Covid with that caused by recent influenza seasons, in a scenario where all the population had been vaccinated.

JEL: I10, I18, H12
Keywords: COVID-19, vaccination, anti-vaxxer, influenza

*This work has been funded by the project PID2020-116242RB-I00 of the Spanish Ministry of Science and Innovation (MCIN/AEI/ 10.13039/501100011033) and by the Generalitat Valenciana grant GVPROMETEO2020-083. José E.Boscá and Javier Ferri are grateful for funding from the Rafael del Pino Foundation and BBVA Research. We thank participants at the XXIX Encuentro de Economía Pública and seminar at the BBVA Research for their helpful comments and suggestions.

†José E. Boscá, Department of Economics, Avda Naranjos s/n, 46022, Valencia, email: jose.e.bosca@uv.es
‡José Cano, Department of Economics, Avda Naranjos s/n, 46022, Valencia, email: jose.cano-leiva.@uv.es
§Javier Ferri, Department of Economics, Avda Naranjos s/n, 46022, Valencia, email: francisco.ferri.@uv.es

1 Introduction

The health crisis triggered by the Covid-19 pandemic has evolved over the two years since the virus was first detected. Various health measures have been taken in an effort to influence and adapt to the course of the pandemic. However, many of the messages conveyed to the public by health authorities, central and regional governments, and the media are still contradictory. One factor that certainly does nothing to reduce uncertainty is the difference between the information that is transmitted in near real time, and the information that is published after undergoing a filtering process. If the information reaching the public on a day-to-day basis differs from the corrected information that ends up being published after passing through certain filters over time (for simplicity's sake, we can call it refined information), agents' reaction would not have been optimal compared to reactions based on the most accurate information. The greater the difference between the two sources of information, the greater the divergence between actual behavior and optimal behavior under refined information. In this paper, we present evidence of notable differences between the information published by the Ministry of Health in its daily bulletins in PDF format, and the refined information published by both the Center for the Coordination of Health Alerts and Emergencies (Centro de Coordinacion de Alertas y Emergencias Sanitarias) and, above all, the National Epidemiological Surveillance Network (RENAVE).

The information content of new infection data has varied over time. Although relevant, the rise in new infections is less of a concern if it is accompanied by fewer hospitalizations or ICU admissions. In this paper, we also document how the relationship between ward and ICU hospitalizations has varied in relation to reported infections, and use this information to provide an overview of the different waves in terms of *infections generating hospital pressure and ICU pressure*.

The sixth wave (the last in Spain to date) has resulted in an exponential increase in the number of infections, unprecedented since the beginning of the pandemic, causing a sharp rise in the number of visits to primary care centers. Said visits are primarily associated with requests for Covid diagnostic tests. From the information on the number of tests performed, in another of the exercises we design a simple arithmetic calculation to put the work overload in primary care centers into context. Specifically, we calculate the increase in the number of health care workers needed to keep the workload per employee constant over time.

Vaccines have been presented as an effective tool in the fight against the disease. However, the lack of information consistent with the official data series in this regard is striking. Using official sources on the evolution of the pandemic, in this note we evaluate the importance of vaccination in reducing infections, hospitalizations, ICU admissions and deaths. We also examine whether there is any relationship between vaccination and the average length of hospital stays. In relation to the effect of vaccines, and under the assumption that the vast majority of those who were not vaccinated when the sixth wave started had made the decision voluntarily (which we can call antivaccine attitude), we evaluate cost in health terms of vaccine hesitancy. Finally, we compare the incidence of Covid-related hospitalizations and deaths in a fully vaccinated population with that of a typical influenza season.

The rest of this document is organized as follows: section (2) presents the discrepancies between the information published daily in the bulletins and the subsequently refined information; section (3) characterizes the waves of infections and hospital pressure during the year 2021; section (4) discusses the changes in the average length of stay in hospital; section (5) calculates the number of infections under the assumption that they would have the same capacity to generate hospital and ICU admissions throughout the year; section (6) provides a simple estimate of the pressure on primary care centers during the sixth wave; section (7) estimates the effect of vaccines on health series and performs different counterfactual exercises; section (8) offers some proposals to reduce the impact of the virus on primary care and hospital centers in the Spanish health system. The last section (9) presents the main conclusions.

2 Covid data in Spain: discrepancies between data sources

The daily update bulletins released by the Center for the Coordination of Health Alerts and Emergencies in PDF format (hereinafter *Bulletins* for short) provide daily information on new infections, hospitalizations and deaths. This is the information that is basically transmitted to the general public through the media. The same agency also produces open data on healthcare capacity (hereinafter *Open Data*) which provides this same information by province, in electronic format and also on a daily basis. This database undergoes modifications over time due to a data cleaning process.

Figure 1 shows the relative difference, in percentages, between the data from the Bulletins and the Open Data, in relation to ward hospitalizations and ICU admissions. In general, the daily data published in the Bulletins overestimate the number of patients admitted to the ward due to Covid-19 by an average of 2 percentage points, although this difference is very volatile, reaching an average difference of more than 4 percentage points in December, with some days exceeding 8 percentage points. Although the differences in reporting relative to ICU admissions are somewhat better, the fact remains that during the month of December the daily bulletins overestimated actual admissions by an average of more than 3 percentage point reaching errors of more than 6 percentage point during the first two weeks of the month. For this reason, whenever possible, we have preferred to use *Open Data* rather than *Bulletins*.

To address some of the aspects dealt with in this study, we rely on age-disaggregated information on confirmed cases reported to the National Epidemiological Surveillance Network (RENAVE, hereinafter). These data provide daily information on the number of infections, hospitalizations, ICU admissions and deaths by sex, province and age group. Unfortunately, there are considerable differences between the Open Data and RENAVE series in terms of the level of data[1]. Figure 2 shows the daily evolution of hospital and ICU admissions with

[1] 1According to RENAVE, these differences are due to the fact that the Open Data come from the

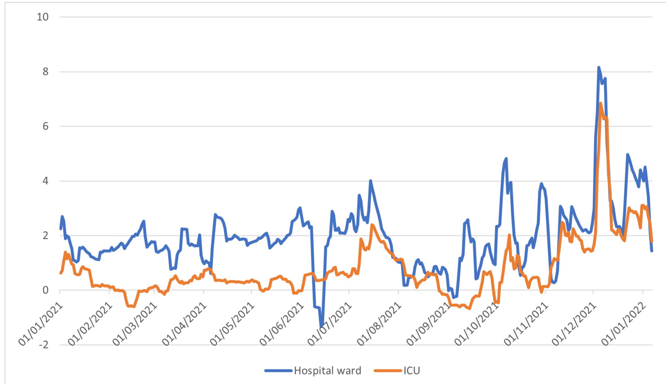


Figure 1: *Differences (in %) between Bulletins and Open Data*

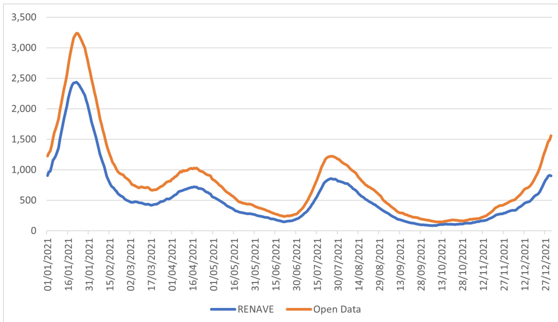
both sources of information (2a and 2b), as well as the differences in percentages between the two (2c). As can be seen, although the profile of the series obtained with data from both sources is very similar, Open Data overestimates the influence of the epidemic on health indicators. In relative terms, these differences, on average during 2021, are 31 percent for hospitalizations and 63 per cent for ICUs.

3 Waves of contagions and hospital pressure

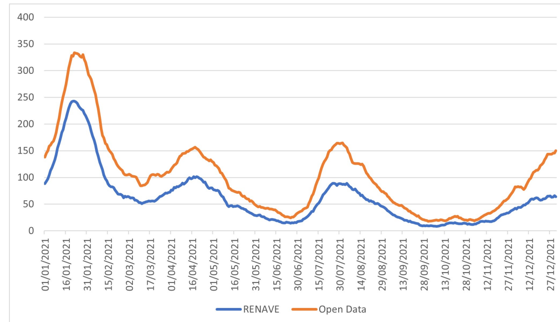
Health data trends during 2021 are displayed in Figure 3. Subplot 3a shows the evolution over time of new daily Covid cases detected during 2021. The year began in the middle of the third wave, which started in early December 2020. There was a very mild fourth wave starting in March. The fifth wave, which was more intense in terms of infections, started at the end of June. The sixth and last wave is unparalleled in terms of both the number of confirmed infections and the growth rate. In the first part of the wave, the rise in infections was basically driven by the *delta* variant of the coronavirus, which was overlapped by the new *omicron* variant, as can be seen in the graph by the acceleration in the rate of infection in the last weeks of the year.

Subplots 3c and 3d show the percentage of beds occupied by Covid-19 patients in hospital wards and ICUs. There is a lag of approximately 10 days between the growth of infections and increasing hospital pressure on the ward, and between 15 and 20 days between infections and pressure on ICUs. On the other hand, as can be seen in the adjusted trend line, the pressure on the wards and on ICUs clearly declined over the course of 2021. At the end of the year, the percentage of occupied ward beds in Spain was approximately 9 percent, much lower than the 22 percent reached in the third wave, with three times more confirmed daily infections than at the peak of the third wave. The aggregate pressure on the ICUs was

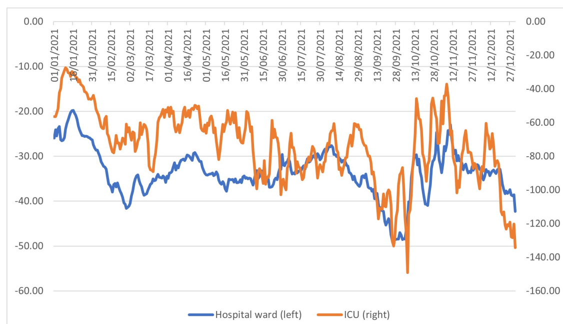
reports provided by each hospital on new admissions to COVID-19 beds, which include people hospitalized with COVID, for COVID, or even with suspected COVID that is not subsequently confirmed. This means that the number is systematically higher than the number of admissions recorded by RENAVE, where in theory the admissions are only for COVID-19, excluding those admitted with COVID-19 and including only confirmed cases.



(a) *Differences in hospitalizations*



(b) *Differences in ICU admissions*



(c) *Differences in hospitalizations and ICU admissions (%)*

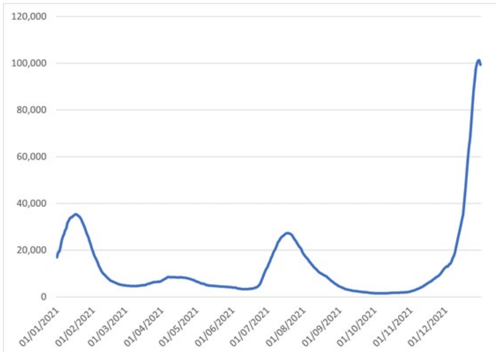
Figure 2: *Differences between Open Data and RENAVE*

around 20 percent, a long way off the 44 percent reached during the third wave, before the population had the opportunity to be fully vaccinated. Not surprisingly, the daily deaths shown in subplot 3b stood at just under 120 people at the end of 2021, a noteworthy figure but six times lower than the peak in daily deaths registered at the beginning of February of that year.[2].

4 The average length of stay in hospital

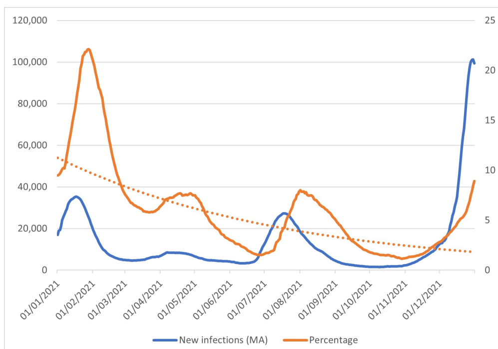
Figure 4 shows the average length of stay in hospital wards and ICUs. ICU admissions are divided, in turn, between patients that required a ventilator and those that did not. To estimate the duration, it has been assumed that the probability of leaving the hospital (or ICU) that existed when the patient was admitted to the hospital (or ICU) will remain constant during their stay. The probability of leaving the care unit (hospital ward or ICU) is then determined on the day of admission by calculating the sum of hospital discharges and deaths divided by the number of persons admitted that day. Under these assumptions, the estimated length of stay is the inverse of this probability.

[2]Daily death data were obtained from the daily PDF bulletins.

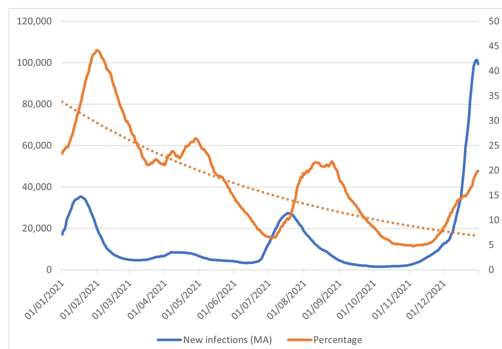


(a) *Infection waves during 2021*

(b) *Daily deaths due to Covid*



(c) *Confirmed infections (left) and percentage of hospital beds occupied by Covid patients (right).*



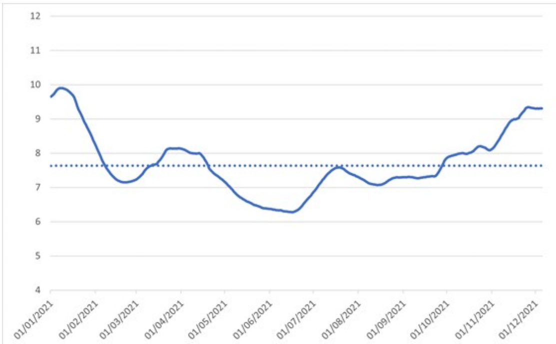
(d) *Confirmed infections (left) and percentage of ICU beds occupied by Covid patients (right).*

Figure 3: *Health data trends during 2021*

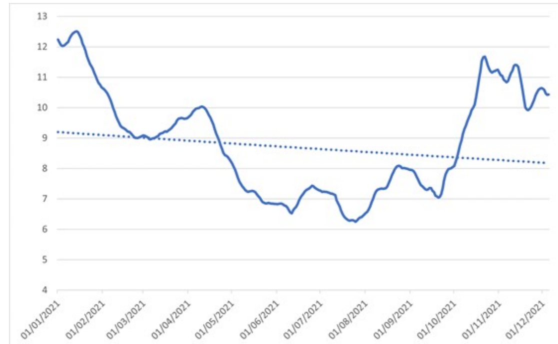
It can be seen that the length of stay in hospital presents notable variation over time. For example, over the course of 2021, the mean length of stay for non-ventilated patients in ICUs (4b) fluctuated between 12 days in the first month of the year, and 7 days in the summer months. The existence of these oscillations in the duration of admission is compatible with a relatively constant trend over the year as a whole (dotted line). According to this trend, the mean length of stay on the hospital ward is just under eight days (4a), only one day less than the mean length of stay in non-ventilator ICUs. However, the ventilator requirement considerably increases the duration of ICU admission to 23 days on average (4c).

5 Homogeneous infection series corrected for disease severity

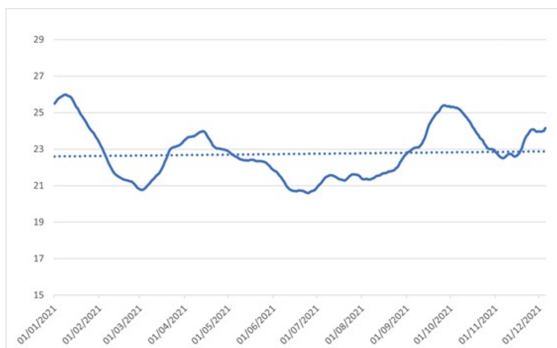
An aspect of great importance, which allows us to characterize the evolution in the severity of the disease, is the potential for those infected with coronavirus in each period to end up requiring hospital care or admission to ICUs (Figure 5). To capture this factor, we divided the



(a) *Average length of hospital admissions (days)*



(b) *Average stay (days) in non-ventilator ICUs*



(c) *Average stay (days) in ventilator ICUs*

Figure 4: Average hospital stays

number of new hospital admissions by the number of new infections 10 days earlier, assuming a lag between the detection of infection and admission to hospital, which is consistent with the evidence presented above (subplot 5a). The exercise is repeated for ICU admissions, but this time assuming a 15-day lag (subplot 5b). In both cases a marked downward trend is observed, with a significant drop during the month of December, in the midst of a wave of *omicron* infections.

The potential of the virus to generate hospital and ICU admissions peaked at the end of March and beginning of April, coinciding with the growing phase of the fourth wave. On the other hand, hospital admissions per infected person, both in wards and ICUs, were at their lowest levels at the end of the year.In fact, hospital admissions (ICU admissions) per 1000 infections fell from 79 (13) persons during the fourth wave to 10 (1.3) persons during the sixth wave, i.e. the virus's ability to generate hospital admissions (ICU admissions) dropped from the year's peak to one eighth (one tenth) of this value.

Having documented this fact, in Figure 6 we examine what the wave profile would have been if infections in the different waves during the year 2021 had generated exactly the same number of ward and ICU admissions per 1000 infected patients.

(a) *Hospital admissions per 1000 cases*

(b) *ICUs admissions per 1000 cases*

Figure 5: *Hospitalization and ICU admission rates*

To this end, we date the onset of each wave and calculate the elapsed time from the onset to the peak of each wave. Let us call that period T_k, where k refers to the wave in question ($k = 3$, 4, 5, 6). Next, we calculate the cumulative number of infections during that time,

$$ZA_{T_k} = \sum_{t \in T_k} Z_t \tag{1}$$

where Z_t refers to new infections in t and ZA_{T_k} to the cumulative cases during the beginning and the peak of wave k.

At the same time, we also calculate the cumulative admissions to the corresponding hospitals,

$$IHA_{T_k} = \sum_{t \in T_k} IH_{t+10} \tag{2}$$

where IH_{t+10} refers to new ward admissions at $t + 10$ and IHA_{T_k} to the cumulative ward admissions during the beginning and the peak of wave k, with a 10-day lag.

The equivalent of the above expression for ICU admissions is as follows:

$$IUA_{T_k} = \sum_{t \in T_k} IU_{t+15} \tag{3}$$

where IU_{t+15} refers to new admissions to ICUs in $t + 15$ and IUA_{T_k} to cumulative ICU admissions during the beginning and peak of the wave k, with a 15-day lag.

From the above expressions we can obtain a factor of average hospital admissions (ward and ICU) per infected patient in wave k

$$f_k^H = \frac{IHA_{T_k}}{ZA_{T_k}} \text{ for } k = 3, 4, 5, 6$$

$$f_k^U = \frac{IUA_{T_k}}{ZA_{T_k}} \text{ for } k = 3, 4, 5, 6$$

Once the factor with the highest value among the above factors has been identified (let us call f_{k*}^H and f_{k*}^U them), correction factors can be computed

$$s_k^H = \frac{f_k^H}{f_{k*}^H} \text{ for } k = 3, 4, 5, 6$$

$$s_k^U = \frac{f_k^U}{f_{k*}^U} \text{ for } k = 3, 4, 5, 6$$

These are used to obtain the Hospital Admissions Generating Cases (HAC) and ICU Admissions Generating Cases (IAC), the results of which are shown in Figure 6. Given that, as we have already seen, the infections that had the greatest capacity to generate admissions to hospitals and ICUs occurred in the fourth wave ($k^* = 4$), it is in the remaining three waves that a flattening of the curve of infections would occur. The orange and green lines in Figure 6 are therefore interpreted as the profile that the waves of infections would have had if the infections in all waves had resulted in as many hospital admissions as in the fourth wave.

HAC and IAC divide by a factor greater than 6 the observed series of infections during the sixth wave. From this perspective, by the end of 2021, the daily HAC would be approximately 16 thousand, and the IAC approximately 13 thousand, still well below the peak of the 22 thousand HAC and IAC of the third wave.

Figure 7 represents an alternative correction in which, instead of using the cumulative number of new admissions, we use the difference during the period T_k between the total number of patients hospitalized at the peak and at the beginning of the wave (with a lag of 10 days in the case of hospitalized patients and 15 days in the case of those admitted to ICUs). In other words, all we do is replace expressions (2) and (3) with the following expressions

$$VH_{T_k} = H_{t_p+10} - H_{t_i+10} \tag{4}$$

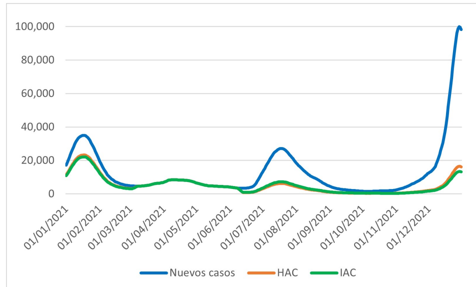


Figure 6: *HAC and IAC*

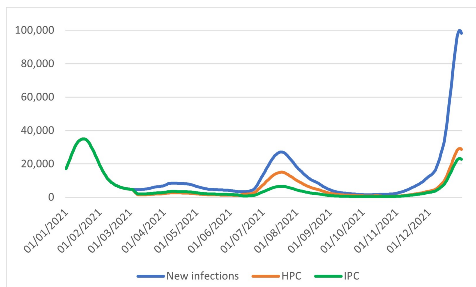


Figure 7: *HPC and IPC*

$$VU_{T_k} = U_{t_p+15} - U_{t_i+15} \tag{5}$$

where H_{t_p+10} and H_{t_i+10} represent, respectively, the hospitalized stock at the peak and at the beginning of the wave (with a lag of 10 days, and VH_{T_k} is the variation in the hospitalized stock during the growth period of wave k infections (with a lag of 10 days). The same interpretation applies to the change in ICU admissions, VU_{T_k}.

This correction better captures the impact of infections on hospital pressure by also considering hospital outflows. For this reason, we call the series corrected in this way Hospital Pressure Generating Cases (*HPC*) and ICU Pressure Generating Cases (*IPC*). As previously indicated, the longest duration of hospital stays on the ward and in ICUs occurred during the third wave, so the infections that were most likely to generate pressure on hospitals and ICUs occurred during this wave ($k^* = 3$).

The conclusions obtained with this correction are quite closely in line with the admission generating cases. Based on the principle of infections generating the same hospital or ICU pressure throughout 2021, the cases detected during the sixth wave should be reduced by a factor of between 3.6 and 4.5. Thus, by the end of 2021, HPC would be approximately 29 thousand, and IPC approximately 23 thousand, still below the peak of the 36 thousand HPC and IPC of the third wave.

6 Pressure on primary care in the sixth wave

The very rapid growth in the number of cases detected during the sixth wave has led to an increase in the number of visits to primary care centers. In this section we provide a metric to quantify the pressure that the sixth wave of infections has placed on these centers. Our approach can be considered a simple back-of-the-envelope calculation, and is based on two principles: *homogeneity* and *prudence*.

In its strictest version, the principle of homogeneity would mean that infections are distributed homogeneously among the districts assigned to the different primary care centers according to the size of the center; applied less strictly, this principle would imply that the human resources of the national health system are allocated efficiently in time and space throughout Spain. The more unequal the health situation, in terms of infection rates, between the different health districts in the country as a whole, and the greater the restrictions on the reallocation of resources between districts, the greater the deviation of the real cost from that reported in our exercise[3].

In accordance with the principle of prudence, we prefer to make assumptions that tend to be at the upper limit of the cost that the sixth wave may impose on primary care centers. For example, we consider that all visits ending with a Covid diagnostic test have taken place in primary care centers and in the public health system. These figures therefore include Covid-19 patients who are diagnosed in public or private hospitals and in private primary care centers.

We also assume that all sick leave management is also done in public sector primary care centers, and that all those diagnosed with the infection will require sick leave. This is again an upper limit to the resource cost of the sixth wave, as in reality many of those infected are children, or the unemployed and inactive population, who do not require sick leave.

On the other hand, our calculations are based on the cautious assumption that the coronavirus epidemic has not affected the number of infections with other diseases. For example, we consider that seasonal influenza infections remain within the normal values of pre-pandemic years, i.e., there is no reduction in the pressure on health centers due to a drop in influenza cases.

The principle of prudence is also applied to the times assigned to the different activities carried out in primary care centers related to new infections. Among these activities we distinguish two groups: those related to the diagnosis of the disease, and those related to the management of *bajas* and *altas*[4].

Table 1 shows the assumptions about the time spent on the different tasks in each of

[3] This statement can be rationalized under the assumption that the central planner prefers a more equitable distribution of the cost of the increase in infection rates in the national territory to a less equitable one.

[4] By means of *bajas* and *altas* we refer to the management of the administrative procedure for certification of sick leave and the end of sick leave or fitness to return to work.

Task	*Time (minutes)*	*Professional*
Diagnosis-related activities		
Making an appointment	5	Receptionist/telephonist
Pre-test diagnosis	10	Physician
Conducting Test	10	Nurse
Processing Test	10	Lab technician
Activities related to *bajas* and *altas*		
Appointment (leave + return)	10	Receptionist/telephonist
Sick leave + return	20	Physician

Table 1: Assumptions about the time required per patient

these two groups of activities, as well as the type of professional involved. From the start of the sixth wave until January 14, 2022, a total of 89 days elapsed and $17,012,516$ Covid-19 diagnostic tests were performed, according to the Ministry of Health's Open Data. Using the assumptions in Table 1, we can transform these tests into minutes involved, minutes into hours, hours into workdays of 7.5 hours per day, and assuming a five-day work week, we obtain the total number of professionals working 37.5 hours per week involved exclusively in Covid-19 diagnosis in primary care centers during the sixth wave.

On the other hand, $3,610,015$ cases were detected in those 89 days. Assuming that all of them require reports certifying the beginning and end of sick leave, and proceeding in a similar way to that described in the previous paragraph, we obtain an estimate of the total number of professionals involved exclusively in the management of paperwork related to Covid-19 sick leave.

The results of the calculations can be found in Table 2. As we have already indicated, we consider these results to be the upper limit. According to our calculations, the total number of staff required to cope with the pressure on primary care centers caused by the sixth wave would be $26,515$ professionals. According to the Main Data of the National Health System published by the Ministry of Health in May 2021 (Ministerio de Sanidad, 2021), there are $13,000$ primary care centers in Spain, so it would be necessary to increase the staff of each center by slightly more than 2 professionals on average to compensate for the excess workload caused by the sixth wave.

Table 3 shows the increase in personnel required by professional category. The Ministry of Health (2021) distinguishes between physicians, nursing staff and other professionals, whose staffing levels are shown in the second column. The category "other professionals" includes receptionists and laboratory technicians. The third column shows the increase in staffing for each professional required to keep the workload constant at the levels prior to the sixth wave. There are three ways to interpret these results. The first is the increase needed in the number of workers in each professional category. Another way to interpret these figures would be

Days elapsed	89	
Tests performed	17,012,516	
Cases detected	3,610,015	
Task	*Increase in the number of professionals*	*Professional*
Diagnosis-related activities		
Making an appointment	3,675	Receptionists
Pre-test diagnosis	7,350	Physicians
Conducting Test	7,350	Nurses
Processing Test	7,350	Technicians
Subtotal	25,729	
Activities related to *bajas* and *altas*		
Appointment (leave + return)	1,262	Receptionists
Sick leave + return	2,524	Physicians
Subtotal	3,786	
Total	29,515	

Table 2: Demands on primary care centers for the sixth wave (staffing requirements)

	Number of current professionals	*Workload (%)*
Diagnosis-related activities		
Physicians	43,000	17.06%
Nurses	39,000	18.85%
Others	34,000	32.43%
Activities related to *bajas* and *altas*		
Physicians	43,000	5.87%
Others	34,000	3.71%

Table 3: Demands on primary care staff for the sixth wave (workload)

in terms of the excess workload imposed on each professional category if no additional staff were hired. A third interpretation would be as a measure of the reduction in care provided to other patients to keep staffing and workloads constant.

According to these results, the sixth wave would have increased the aggregate workload of the Other Professional category, which includes receptionists/phone operators and also laboratory technicians, by more than 36 per cent. The aggregate workload of physicians would have increased by about 24 percent, of which about 6 percent is due to managing the certification of the beginning and end of sick leaves. For nurses, the increase in workload would have been about 19 percent.

Underlying the above calculations is the implicit assumption that the workload has been evenly distributed since the beginning of the sixth wave, although the cumulative workload over the intervening time is consistent with the cumulative number of cases detected and tests performed. In reality, infections have accelerated over time, so the workload would be higher in the last month and lower at the beginning of the wave. In fact, if we repeat the calculations only for the cases detected between December 15 and January 14 (the last month with available data), we obtain an increase (with respect to the pre-pandemic situation) of $46,008$ professionals needed to cope with the extra workload in the last month 39% more physicians, 27% more nurses and 55% more other professionals, including laboratory technicians). As we have indicated, we consider these figures to be the upper limit of the workload increase.

7 The health effect of vaccination

The scientific community is in no doubt as to the effectiveness of vaccination in reducing transmission and preventing the development of severe disease after Covid-19 infection. To date, however, there is no empirical evidence on the extent to which vaccination has contributed to effectively reducing the incidence of Covid-19 in Spain. In this section we provide an estimate of the effect of vaccination on infections, hospitalizations, ICU admissions and deaths during 2021.

The specific data used to carry out the simulations in this section come from different official sources. First, in Table 10 of the update no. 529 of the Bulletin of December 23, 2021 from the Ministry of Health, information appeared (sporadically) on the number of infections, hospitalizations and deaths, distinguishing by age group and according to whether the affected person was vaccinated or not. From this table we can define the average incidence of new infections (per 100,000 inhabitants) among the population with a vaccination status d in age group e (let us call it $z^{d,e}$) during a window $\widetilde{s}$ of s weeks for which we have information (i.e. the 8 weeks between October 18, 2021 and December, 12, 2021). The d index can refer to the fully vaccinated population (vc), to the population that has not received any Covid vaccine (vn), or to the population that has received a dose of a vaccine, but is not fully vaccinated (vi). That is to say $d = \{vc, vn, vi\}$. Furthermore, the index e divides the population into groups aged from 12 to 29 ($e = 1$); from 30 to 59 ($e = 2$); from 60 to 79 ($e = 3$); and older than 80 ($e = 4$); thus, $e = \{1, 2, 3, 4\}$.

Secondly, data from RENAVE (2021) are used[5]. Information from a one-off table from the Ministry of Health (2021b) has also been used, as well as population data by age from the National Institute of Statistics, INE (2021). The data are worked in terms of weekly flows, from the first week of April 2021, since this is when data on the vaccinated population

[5]This is the only source that makes it possible to distinguish influence on health indicators by age. The comparison between this data file and the Open Data from the Ministry of Health shows that the figure for total cumulative hospitalizations is about 30 percent lower in the RENAVE data than in the Open Data, and total ICU admissions is about 60 percent lower.

according to vaccination status and by age groups become available

The basic assumption we use is that the average incidence rates of the vaccinated relative to those of the unvaccinated and the partially vaccinated remain constant during 2021. That is, if we define the mean relative incidence of the vaccinated to the unvaccinated (x), and to the partially vaccinated (y), by age group, this relative incidence is assumed to be unchanged over time.

$$x_z^e = \frac{z^{vc,e}}{z^{vn,e}}; \quad x_h^e = \frac{h^{vc,e}}{h^{vn,e}}; \quad x_u^e = \frac{u^{vc,e}}{u^{vn,e}}; \quad x_f^e = \frac{f^{vc,e}}{f^{vn,e}} \tag{6}$$

$$y_z^e = \frac{z^{vc,e}}{z^{vi,e}}; \quad y_h^e = \frac{h^{vc,e}}{h^{vi,e}}; \quad y_u^e = \frac{u^{vc,e}}{u^{vi,e}}; \quad y_f^e = \frac{f^{vc,e}}{f^{vi,e}} \tag{7}$$

where z, h, u and f refer to infections, hospitalizations, ICU admissions and deaths. The longer the duration of vaccine effectiveness, or the sooner the booster dose is given, the more easily this assumption is satisfied. In our opinion, in most cases of complete vaccination, the time elapsed since the last vaccination guarantees the complete effectiveness of the vaccine and, therefore, we believe that the assumption is very reasonable.

Under this assumption, we can obtain an estimate of the incidence of infection among the fully vaccinated in age group e, which is unobserved over time, from the expression (see *Appendix A* for derivation)

$$\widehat{r}_j^{vc,e} = \frac{r_j^e}{\left(\frac{P_j^{vc,e}}{P^e} + \frac{1}{x_r^e}\frac{P_j^{vn,e}}{P^e} + \frac{1}{y_r^e}\frac{P_j^{vi,e}}{P^e}\right)} \tag{8}$$

where subscript j refers to week j of the year 2021, r can refer to z, h, u or f; P^e is the Spanish population in age group e; and $P_j^{vc,e}$, $P_j^{vn,e}$, $P_j^{vi,e}$ represent, respectively, the vaccinated, unvaccinated and partially vaccinated population in age group e in week j.

Thus, to estimate the incidence (infections, hospitalizations, ICUs, deaths) among the fully vaccinated in age group e in week j of the year 2021, we need the observed incidence in that age group, the relative incidences (x, y) calculated as above, and the percentages of the fully vaccinated, unvaccinated and partially vaccinated population by age group.

Note that the difference in week j between the observed incidence by age group, r_j^e, and that estimated for the fully vaccinated by age group, $\widehat{r}_j^{vc,e}$, will be greater the smaller x_z^e (the average relative incidence of the vaccinated relative to the unvaccinated) and the greater the proportion $\frac{P_j^{vn,e}}{P^e}$ of the unvaccinated in the population.

Our procedure allows us to perform different counterfactual exercises using estimation (8). Details on the procedures used can be found in *Appendix A*, while *Appendix* explains some corrections made to the original data in order to make them consistent with the population information published by INE. Some of the counterfactual simulations performed are described below.

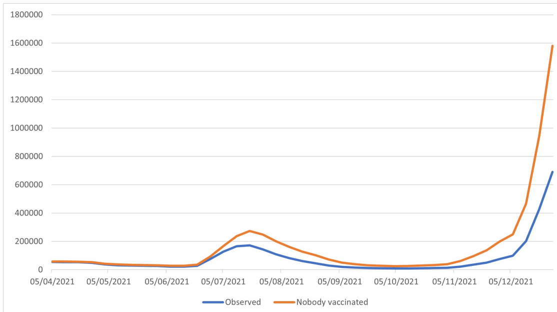
7.1 Counterfactual 1: Vaccine not available during 2021

This first exercise seeks to describe the situation in a scenario in which the vaccine had not reached the Spanish population. Given the same social behavior as that observed, how would the health indicators have evolved? To carry out this exercise we first simulate the evolution of the series of infections, hospital admissions, ICU admissions and deaths, assuming that the vaccines had not existed, and subtract those observed. This difference will give us the effect of vaccination in Spain *for the same social behavior* in the counterfactual scenario as in the observed scenario (i.e. the results would be different if, in the absence of vaccination, the competent authorities had restricted social interactions beyond the restrictions in effect during 2021).

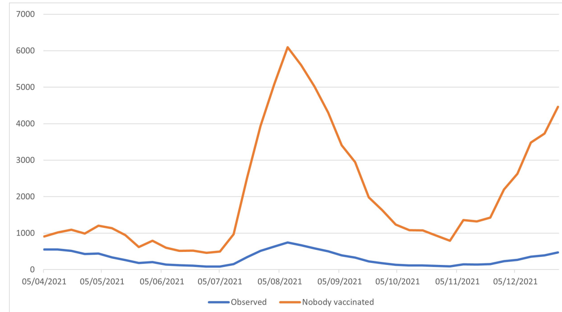
Figure 8 shows the weekly evolution of the flows of infections, hospitalized patients, admissions to ICUs and deaths. The blue line shows the observed evolution and therefore incorporates the effect of the vaccines as the different age groups of the population gain access to them. The orange line represents the evolution of the series in the scenario in which there was no vaccine available during 2021. A marked vaccine effect is observed in all flows, although less so in the infections than in the series capturing the severity of infections. The decoupling since the end of the summer between the observed series and the simulated series in the absence of vaccines is very striking.

The cumulative differences over the year between the two series are shown in Table 4, which divides the population into four age groups. To help interpret this table, let us look at the case of infections. The first row shows the cumulative differences between what would have happened to infections in the absence of vaccination and the actual observed behavior of infections. For example, in the absence of vaccination, cumulative infections would have increased by $3,048,407$ from the first week of April. This increase in infections would naturally lead to immunization of part of the population, so the values shown in Table 4 should be interpreted as an upper limit, although fairly close to the real one. As we can see, the increase in the number of infections would have had a particularly marked effect on the 60-79 age group. The values in the second row represent the increase in percentage terms with respect to the cumulative observed values from April for each age group. In the 80+ age group, infections would have increased by 388 percent. The third row gives an idea of the relative importance of the age group in the total increase in infections in the absence of vaccination. For example, the increase in infections in those under 30 years of age would have accounted for 15 percent of the total increase.

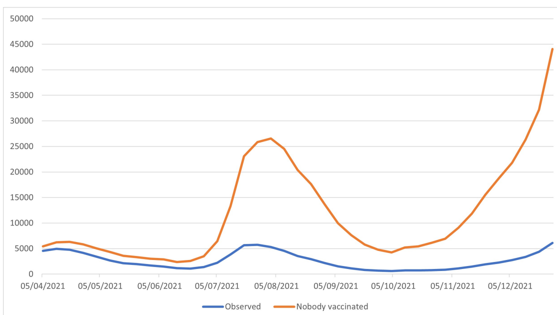
If we now look at the deaths section, we see that if there were no vaccines, Covid-19 deaths would have increased by almost 571 percent . It is interesting to note that those under 60 years of age would account for only 2.1 percent of the total increase, while those over 80 years of age would have been hit the hardest



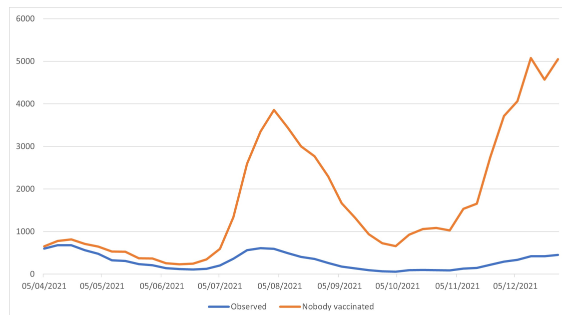
(a) *Weekly cases*



(b) *Weekly deaths*



(c) *Weekly Hospital Admissions*



(d) *Weekly ICU Admissions*

Figure 8: *Health series: observed and counterfactual for non-vaccines*

Overall, the results indicate that the relative benefit of vaccination increases with the severity of the indicator, being greater for deaths and ICU admissions and less notable for infections and hospitalizations.

7.2 Counterfactual 2: the influence of anti-vaxxers on the development of the sixth wave

What would have happened if 100 percent of the population over 12 years of age had been vaccinated when the sixth wave began? By means of this counterfactual exercise we try to establish the influence of the vaccine-hesitant population on the evolution of health indicators during the sixth wave. The starting assumption is that, in the first week of November, the entire population over 12 years of age had been offered all doses of the vaccine, so that the unvaccinated could be considered anti-vaxxers. The difference between our counterfactual, in which we assume that the entire population is vaccinated at the beginning of the sixth wave, and the health data actually observed will give us the influence of vaccine hesitancy on the cumulative values of health indicators in the sixth wave until the end of December.

The results in Table 5 suggest that vaccine-hesitant people during the sixth wave through the end of December would have increased infections by 14%, hospitalizations by 44% (79%

	10 to 29	30 to 59	60 to 79	Older than 80	Total
			Cases		
Difference from observed	596,001	784,586	1,386,005	281,815	3,048,407
Difference from observed (%)	52	50	388	335	96
Percentage of total difference (%)	20	26	45	9	100
			Hospitalizations		
Difference from observed	5,730	54,713	198.836	99,624	358,904
Difference from observed (%)	70	132	589	513	349
Percentage of total difference (%)	2	15	55	28	100
			ICU admissions		
Difference from observed	254	8,927	43,315	3,305	55,801
Difference from observed (%)	48	172	762	900	474
Percentage of total difference (%)	0,5	16	77.5	6	100
			Deaths		
Difference from observed	95	1,395	24,996	42,037	68,523
Difference from observed (%)	197	126	562	656	571
Percentage of total difference (%)	0.1	2	36.4	61.5	100

Table 4: Effect of vaccination from April 1, 2021

for the 30-59 age group), ICU admissions by 78% (143% for the 30-59 age group), and deaths by 32% (53% in the 30-59 age group).

7.3 Counterfactual 3: health indicators with the entire population vaccinated; comparison with the incidence of influenza

In this third exercise, we raise the question of what would have happened to the health indicators if the entire population had been vaccinated from the moment we have data on vaccination by age and vaccination pattern, that is, from the first week of April. We then compare the cumulative values from April to December of infections, hospital admissions to the ward, ICUs and deaths, under the assumption that 100 percent of the population would have been vaccinated in that period, with the data on the incidence of influenza in the 2017-2018, 2018-2019 and 2019-2020 seasons. We also offer, for reference, the actual Covid data reported by RENAVE for the entire year of 2021.

The objective of this exercise is to illustrate the impact of the disease on the health of the population and on the health system, in a simulated scenario in which the entire population had been vaccinated against Covid. The results can be found in Table 6. The first column shows the observed results for total infections, hospitalizations, ICU admissions, and deaths from April 2021. The second column shows the simulated aggregates from the first week

	10 to 29	30 to 59	60 to 79	Older than 80	Total
			Cases		
Difference (Observed - All vaccinated)	97,774	80,671	22,521	3,308	204,274
Difference (Observed - All vaccinated) (%)	26	10	13	10	14
Percentage of total difference (%)	48	39	11	2	100
			Hospitalizations		
Difference (Observed - All vaccinated)	731	3,315	2,489	840	7.376
Difference (Observed - All vaccinated) (%)	124	79	34	17	44
Percentage of total difference (%)	10	45	34	11	100
			ICUs Admissions		
Difference (Observed - All vaccinated)	25	507	536	33	1.101
Difference (Observed - All vaccinated) (%)	80	143	57	40	78
Percentage of total difference (%)	2	46	49	3	100
			Deaths		
Difference (Observed - All vaccinated)	9	65	215	257	546
Difference (Observed - All vaccinated) (%)	-	53	38	25	32
Percentage of total difference (%) (%)	2	12	39	47	100

Table 5: Effect of the unvaccinated population in the sixth wave

of April, under the assumption that the entire Spanish population over 12 years of age had been fully vaccinated against coronavirus. The following three columns provide information on the incidence of influenza according to the Spanish Influenza Surveillance System of the National Epidemiological Surveillance Network, for the 2017/18, 2018/19 and 2019/20 seasons. Note that we are comparing complete influenza seasons, with only the last three quarters of 2021 for Covid.

In any case, comparing columns two and three does not reveal particularly unfavorable differences in the damage caused by Covid-19, under the assumption of the entire population being vaccinated, with respect to a seasonal flu season that was rated as having moderate/high activity. In fact, mortality caused by 2017/18 flu would far exceed that caused by Covid-19 during 2021. Infections are not comparable as influenza diagnostic tests are not part of the routine screening procedure. We must introduce, however, two important considerations: (a) the 2021 season for Covid-19 is incomplete, as information for the first quarter is missing; (b) not all the population at risk is vaccinated against influenza.

In any case, the results introduce a certain dose of optimism about the possibility that, in the near future, the influence of coronavirus on the health system, with a population suitably aware of the need for vaccination, will be similar to that of influenza. Covid-19 is a new disease caused by a new coronavirus. The population and those responsible for the health system must adapt to this structural change and normalize it. As with influenza, we will never be sure of the degree of activity of the new variants that may appear. Fortunately,

	Covid (obs) Apr-Dec 21	Covid (vac) Apr-Dec 21	Flu (17/18)	Flu (18/19)	Flu (19/20)
Cases	3,169,872	2,319,807	70,000	490,000	619,000
Hospitals	102,847	44,763	52,000	35,300	27,700
ICUs	11,764	3,443	3,000	2,500	1,800
Deaths	12,008	6,601	15,000	6,300	3,900

Table 6: Influence on health indicators with complete vaccination

vaccination is our ally in this fight: as we have shown in these exercises, it has reduced the health impact of the epidemic to the level of that of a bad flu season.

8 Proposals for reducing the health and economic impact of coronavirus.

From the reading of our calculations and estimates in the previous pages, there is a first conclusion that is very important to highlight, even if it is already widely known: the vaccines have been effective—very effective in fact—and have significantly changed the course of the pandemic. With all due caution, it can be noted that the number of people who develop severe disease and, therefore, are admitted to ICUs or eventually die, has decreased drastically as a percentage of total infections. Our estimates should be taken with some caution because, although they reflect the situation in 2021, they only partially reflect the incidence of the *omicron* variant of the virus. In any case, from what we know so far, it does seem that this variant, although much more contagious, is less likely to cause severe disease (by itself or in conjunction with the mass vaccination program in our country), as can be seen in our data from the sixth wave.

If, as it seems, we are condemned to a near future in which the Covid-19 virus does not disappear, but has a recurrent seasonal character, what can be done to reduce the economic and health impact of Covid-19 when new waves of the virus appear?

First, health authorities should develop contingency plans to be able to efficiently reinforce the primary care system, the hospital system and intensive care. In this regard, correctly assessing the workload increases that occur in situations of pandemic stress is of paramount importance; based on this information, extra staff can be hired to ensure that other medical needs are met. In such situations of pandemic stress, permanent or backup healthcare personnel should be able to be temporarily transferred from their usual work stations to others that are saturated. It is a fact that on many occasions the incidence of the virus is very unequal between health districts (or municipalities, provinces, or regions) and there should be no reason why human resources cannot be mobile. Indiscriminately increasing the number of permanent positions does not seem to be the most sensible option if there are going to be peaks of excess work in the future (as there always have been; for example, in flu seasons).

Secondly, it seems sensible to intensify incentives of all kinds to encourage the unvaccinated to get vaccinated. Without having to curtail individual freedom, health passports (or daily tests) should be required for access to all kinds of public and private events in enclosed spaces or mass events in open spaces. The idea is that unvaccinated people would find it more difficult to access leisure and recreational activities and this would persuade them to get vaccinated. These measures can be justified by the social cost the unvaccinated generate.

Third, rapid and efficient procedures for antigen or PCR testing should be established. The collapse of primary care is due in part to medical procedures that are neither technically demanding nor particularly time-consuming, such as diagnosis of the disease via tests or certifying the beginning and end of sick leave. In large cities, test centers could be set up in different parts of the city, independent of the primary care centers. With very few medical, health and administrative personnel they can manage testing and sending the results to the patients' cell phones. In the same way, the results could be communicated to the administration in charge of sick leave, establishing streamlined administrative procedures to manage the sick leave and subsequent return-to-work authorizations. For example, an employee could request sick leave as soon as a patient has a confirmed positive test. The sick leave could then be automatically ended within the timeframe decreed by the experts (5, 7 or 10 days), unless the patient required additional care, in which case the regular physician could continue with treatment.

Finally, another aspect that would alleviate the burden on public health care in situations of pandemic stress would be to strengthen coordination and cooperation with the private sector. The private healthcare sector can assist in healthcare management in these situations by putting in place contingency plans that allow for the establishment of agreements. The possibility of referring patients, tests, hospitalizations or ICU stays to private healthcare could in many cases help to relieve the strain on public healthcare.

9 Conclusions

This paper provides an overview of the Covid-related health situation in Spain during 2021, showing how far we have come, and indicating what we can expect in the future. The overexposure of experts, journalists, talk-show hosts, or politicians giving their opinion on the pandemic tends to generate contradictory or confusing messages. Our aim in this study has been to answer, with the support of official data and relying on cautious assumptions, relevant questions that frequently appear in the public debate: How has the pressure on hospitals and ICUs evolved over time? Has the fatality rate of the virus changed? What about the length of hospital stays? What would the waves of contagion look like if we assumed the same probability over time that an infected person would develop severe disease? How much strain is the sixth wave putting on primary care? How many infections, hospitalizations and deaths has vaccination prevented? What is the cost of vaccine hesitancy in terms of hospitalizations and deaths during the sixth wave? How different would a scenario of widespread vaccination against Covid-19 be from a standard flu season?

By way of summary, our answers to the above questions are as follows;

1. There is a clear downward trend in the pressure on hospital beds on the ward and in ICUs.

2. The fatality rate of the virus has been drastically reduced during 2021.

3. Conditional on admission to a hospital (ward or ICU) there is no clear trend change in the duration of hospitalizations.

4. The ability of the virus to end up generating severe complications has been declining over time, so if we were to correct for this fact, in the sixth wave at the end of December 2021 we would have been below the peak of the third wave in early 2021.

5. To maintain the workload of primary care staff at pre-pandemic levels, from the beginning of the sixth wave until January 14, 2022, the staff of each primary care center would have had to be increased by more than 2 employees on average. During this period, the aggregate workload of the Other professionals category, which includes receptionists/telephonists and also laboratory technicians, would also have increased by more than 36%. The aggregate workload of physicians would have increased by about 24%, of which about 6 percentage points is due to the management of medical leaves procedures. For nurses, the increase in workload would have been 19%. We consider these estimates to be a comfortably upper limit.

6. The positive impact of vaccination on health indicators has been extraordinary. For example, the non-availability of vaccines would have increased Covid-19 deaths by 571 percent compared to observed deaths, although those under 60 years of age would account for only 2.1 percent of the total increase, while those over 80 years of age would have been hardest hit. Moreover, the relative benefit of vaccination increases with the severity of the indicator: greater for deaths and ICU admissions, less for infections and hospitalizations.

7. The anti-vaxxers are responsible for a 44% increase in hospitalizations during the sixth wave (79% for the 30-59 age group), a 78% increase in ICU admissions (143% for the 30-59 age group), and a 32% increase in deaths (53% in the 30-59 age group).

8. Under the assumption that the entire population had been vaccinated with the Covid-19 vaccine from the first week of April, the impact of the disease in terms of hospitalizations and deaths during 2021 would have been above a low activity flu season, but in line with a moderate/high activity flu season.

Some of the above results are based on assumptions, which are always made explicit. As we have indicated, our intention is to contribute to the debate with evidence based on public information, and with replicable exercises.

References

1. Centro de Coordinación de Alertas y Emergencias Sanitarias (2021a): .*Actualizaciones diarias. Enfermedad por el coronavirus* (COVID-19). N^o 282 a 533.

2. Centro de Coordinación de Alertas y Emergencias Sanitarias (2021b): *Datos abiertos de la capacidad asistencial.*

3. INE (2021): Población por edad (grupos quinquenales).

4. Ministerio de Sanidad (2021): Principales Datos del Sistema Nacional de Salud publicado por el Ministerio de Sanidad en mayo de 2021.

5. Ministerio de Sanidad (2021b): Informe de Actividad Diario de Gestión Integral de Vacunación frente al COVID-19 en España.

6. Red Nacional de Vigilancia Epidemiológica (2021): Datos notificados a la RENAVE a través de SiVieS. Datos abiertos.

7. Red Nacional de Vigilancia Epidemiológica (2021b): El Sistema de Vigilancia de la Gripe en España.

A A method for estimating Covid-19 severity as a function of vaccination.

Let us consider the week as the unit of time reference in our analysis.

Let $P_j^{d,e}$ be the stock of population that during week j has a vaccination status d, and belongs to age group e. The index d can refer to the population that is fully vaccinated (vc), to the population that has not received any coronavirus vaccination (vn), or to the population that has received a dose of the vaccine, but is not fully vaccinated (vi). That is, $d = \{vc, vn, vi\}$. On the other hand, the index e divides the population into groups aged from 12 to 29 years ($e = 1$); from 30 to 59 years ($e = 2$); from 60 to 79 years ($e = 3$); and older than 80 years ($e = 4$); thus,$e = \{1, 2, 3, 4\}$.

Total population in a population group d is defined as

$$P_j^d = \sum_e P_j^{d,e}$$

Total population in an age group e

$$P_j^e = \sum_d P_j^{d,e}$$

Total population over 12 years of age

$$P = \sum_d \sum_e P_j^{d,e}$$

which, for practical purposes, we will assume to be constant for $\forall j$.

Let us call $Z_j^{d,e}$ the number of *new cases* during week j of the pandemic, among the population group with vaccination status d, and belonging to age group e.

In week j, the total number of new cases in age group e is the sum of fully vaccinated, unvaccinated and partially vaccinated.

$$Z_j^e = \sum_d Z_j^{d,e} \tag{9}$$

And the total number of new cases in the population as a whole

$$Z_j = \sum_d \sum_e Z_j^{d,e} \tag{10}$$

Similarly, we can express the weekly total of *new hospital admissions* (H), *ICU* admissions (U), and *deaths* (F) of age group e in week j as:

$$H_j^e = \sum_d H_j^{d,e} \tag{11}$$

$$U_j^e = \sum_d U_j^{d,e} \tag{12}$$

$$F_j^e = \sum_d U_j^{d,e} \tag{13}$$

And the total number of new hospital admissions, ICU admissions and deaths in the population as a whole

$$H_j = \sum_d \sum_e H_j^{d,e} \tag{14}$$

$$U_j = \sum_d \sum_e U_j^{d,e} \tag{15}$$

$$F_j = \sum_d \sum_e U_j^{d,e} \tag{16}$$

The incidence rate of new cases of *infections* (per 100,000 population) in any given week j among population d and age group e is defined as

$$z_j^{d,e} = 100.000 \frac{Z_j^{d,e}}{P_j^{d,e}} \tag{17}$$

The overall incidence in week j by age group e is then

$$z_j^e = 100.000 \frac{\sum_d Z_j^{d,e}}{\sum_d P_j^{d,e}} = 100.000 \frac{Z_j^e}{P_j^e} \tag{18}$$

where the population in age group e will be assumed constant for $\forall j$, $P_j^e = P^e$

The overall incidence for the population as a whole is

$$z_j = 100.000 \frac{\sum_d \sum_e Z_j^{d,e}}{\sum_d \sum_e P_j^{d,e}} = 100.000 \frac{Z_j}{P} \tag{19}$$

The same definitions can be used to obtain incidence rates of new *hospitalizations, ICUs* and *deaths* in any given week j

$$h_j^{d,e} = 100.000 \frac{H_j^{d,e}}{P_j^{d,e}} \tag{20}$$

$$h_j^e = 100.000 \frac{H_j^e}{P^e} \tag{21}$$

$$h_j = 100.000 \frac{H_j}{P} \tag{22}$$

$$u_j^{d,e} = 100.000\frac{U_j^{d,e}}{P_j^{d,e}} \tag{23}$$

$$u_j^e = 100.000\frac{U_j^e}{P^e} \tag{24}$$

$$u_j = 100.000\frac{U_j}{P} \tag{25}$$

$$f_j^{d,e} = 100.000\frac{F_j^{d,e}}{P_j^{d,e}} \tag{26}$$

$$f_j^e = 100.000\frac{F_j^e}{P^e} \tag{27}$$

$$f_j = 100.000\frac{F_j}{P} \tag{28}$$

Define $z^{d,e}$ the average incidence of new cases of infection (per 100,000 population) among population d and age group e during a window $\widetilde{s}$ of s weeks for which information is available (i.e., the 8 weeks between 10/18/2021 and 12/12/2021).

$$z^{d,e} = 100.000\frac{\sum_{j\in\widetilde{s}}\left(\frac{Z_j^{d,e}}{P_j^{d,e}}\right)}{s} \tag{29}$$

Similarly, $h^{d,e}$ can be defined as the average incidence of new hospitalized cases among the population d of age group e, during the same window of weeks.

$$h^{d,e} = 100.000\frac{\sum_{j\in\widetilde{s}}\left(\frac{H_j^{d,e}}{P_j^{d,e}}\right)}{s} \tag{30}$$

The weekly incidence in ICUs

$$u^{d,e} = 100.000\frac{\sum_{j\in\widetilde{s}}\left(\frac{U_j^{d,e}}{P_j^{d,e}}\right)}{s} \tag{31}$$

And the weekly incidence of deaths

$$f^{d,e} = 100.000\frac{\sum_{j\in\widetilde{s}}\left(\frac{F_j^{d,e}}{P_j^{d,e}}\right)}{s} \tag{32}$$

From official information published by the Center for the Coordination of Health Alerts and Emergencies of the Ministry of Health, the above average incidence rates can be obtained for $d = \{vc, vn\}$. No information is provided, however, on the average weekly incidence for $d = vi$.

We will assume that the incidence of the partially vaccinated population is somewhere in between that of the fully vaccinated population and the unvaccinated population, depending on a parameter λ that captures the closeness of the incidence of this population to that of the vaccinated population. We will consider the case of 0.5, but results are very robust to other likely values of the parameter. Thus

$$z^{vi,e} = \lambda z^{vc,e} + (1-\lambda) z^{vn,e} \tag{33}$$

We can now define the average relative incidence of those vaccinated with respect to those not vaccinated (x) and those partially vaccinated (y).

$$x^e_z = \frac{z^{vc,e}}{z^{vn,e}}; \quad x^e_h = \frac{h^{vc,e}}{h^{vn,e}}; \quad x^e_u = \frac{u^{vc,e}}{u^{vn,e}}; \quad x^e_f = \frac{f^{vc,e}}{f^{vn,e}} \tag{34}$$

$$y^e_z = \frac{z^{vc,e}}{z^{vi,e}}; \quad y^e_h = \frac{h^{vc,e}}{h^{vi,e}}; \quad y^e_u = \frac{u^{vc,e}}{u^{vi,e}}; \quad y^e_f = \frac{f^{vc,e}}{f^{vi,e}} \tag{35}$$

Equation (9) of *observed* infections in a given week j in age group e can be rewritten as

$$Z^e_j = Z^{vc,e}_j + Z^{vn,e}_j + Z^{vi,e}_j = \frac{z^{vc,e}_j P^{vc,e}_j + z^{vn,e}_j P^{vn,e}_j + z^{vi,e}_j P^{vi,e}_j}{100.000} = \frac{z^e_j P^e}{100.000} \tag{36}$$

this is the observed infection information over time by age group. z^e_j is the observed infection incidence rate in week j for age group e, and P^e_j is the population in age group e .

From (36) and equations (34) and (35) we can obtain an estimate of the incidence of infections among fully vaccinated people in the age group e ($z^{vc,e}_j$) which is unobserved over time.

$$\frac{z^{vc,e}_j P^{vc,e}_j}{100.000} + \frac{\frac{z^{vc,e}_j}{x^e_z} P^{vn,e}_j}{100.000} + \frac{\frac{z^{vc,e}_j}{y^e_z} P^{vi,e}_j}{100.000} = \frac{z^e_j P^e}{100.000}$$

$$z^{vc,e}_j \left(\frac{P^{vc,e}_j}{P^e} + \frac{1}{x^e_z}\frac{P^{vn,e}_j}{P^e} + \frac{1}{y^e_z}\frac{P^{vi,e}_j}{P^e} \right) = z^e_j$$

$$\widehat{z}^{vc,e}_j = \frac{z^e_j}{\left(\frac{P^{vc,e}_j}{P^e} + \frac{1}{x^e_z}\frac{P^{vn,e}_j}{P^e} + \frac{1}{y^e_z}\frac{P^{vi,e}_j}{P^e} \right)} \tag{37}$$

To estimate the incidence among the fully vaccinated in age group e, we need the observed incidence in that age group, the relative incidences calculated in Table 1, and the percentages of the fully vaccinated, unvaccinated, and partially vaccinated population.

Note that the difference in week j between the observed incidence by age group and that estimated for the fully vaccinated will be greater the smaller x^e_z (the average relative incidence of the vaccinated relative to the unvaccinated) and the greater the proportion $\frac{P_j^{vn,e}}{P^e}$ of the unvaccinated in the population.

Now we can obtain the estimated infections in week j of age group e under the assumption that the entire population of age e was vaccinated. To do this we change the incidence rate of the unvaccinated and the partially vaccinated to the incidence rate of the fully vaccinated.

$$\widehat{Z}_j^{vc,e} = \frac{\widehat{z}_j^{vc,e} P_j^{vc,e} + \widehat{z}_j^{vc,e} P_j^{vn,e} + \widehat{z}_j^{vc,e} P_j^{vi,e}}{100.000} = \frac{\widehat{z}_j^{vc,e} P^e}{100.000} \tag{38}$$

From expression (38) we can obtain the total number of infected people for the whole population under the assumption that the whole population was fully vaccinated.

$$\widehat{Z}_j^{vc} = \sum_e \widehat{Z}_j^{vc,e} \tag{39}$$

The difference over a period of time between observed infections and those that would have been observed if the population over 12 years of age had been fully vaccinated would be

$$\sum_j Z_j - \sum_j \widehat{Z}_j^{vc} \tag{40}$$

If we want to calculate the difference in a specific period, e.g. the sixth wave

$$\sum_{j \in 6^{th} wave} Z_j - \sum_{j \in 6^{th} wave} \widehat{Z}_j^{vc} \tag{41}$$

Although other types of simulations could also be run. For example, we could calculate the number of infections/hospitalizations/UCIs/deaths if those older than 30 had been fully vaccinated and those younger than 30 had not been vaccinated.

$$\widehat{Z}_j^{30} = \frac{\widehat{z}_j^{vn,1} P^1 + P^e \sum_{e=2}^{4} \widehat{z}_j^{vc,e}}{100.000} = \frac{\frac{\widehat{z}_j^{vc,1}}{x_z^1} P^1 + P^e \sum_{e=2}^{4} \widehat{z}_j^{vc,e}}{100.000} \tag{42}$$

An interesting scenario is that of complete absence of vaccination in the population as a whole, which can be calculated as follows

$$\widehat{Z}_j^{vn} = \frac{P^e \sum_e \widehat{z}_j^{vn,e}}{100.000} = \frac{P^e \sum_e \frac{\widehat{z}_j^{vc,e}}{x_z^e}}{100.000} \tag{43}$$

The effect of the vaccines that have been used so far on infection could be obtained through

$$\sum_j \widehat{Z}_j^{vn} - \sum_j \widehat{Z}_j$$

which would give us the reduction in cases due to the vaccines given during a period of time.

Exactly the same type of experiment could be done for *hospitalization*, *ICU admission* and *deaths*. In this case, the equivalent expressions to (37) are as follows

$$\widehat{h}_j^{vc,e} = \frac{h_j^e}{\left(\frac{P_j^{vc,e}}{P^e} + \frac{1}{x_h^e}\frac{P_j^{vn,e}}{P^e} + \frac{1}{y_h^e}\frac{P_j^{vi,e}}{P^e}\right)} \tag{44}$$

$$\widehat{u}_j^{vc,e} = \frac{u_j^e}{\left(\frac{P_j^{vc,e}}{P^e} + \frac{1}{x_u^e}\frac{P_j^{vn,e}}{P^e} + \frac{1}{y_u^e}\frac{P_j^{vi,e}}{P^e}\right)} \tag{45}$$

$$\widehat{f}_j^{vc,e} = \frac{f_j^e}{\left(\frac{P_j^{vc,e}}{P^e} + \frac{1}{x_f^e}\frac{P_j^{vn,e}}{P^e} + \frac{1}{y_f^e}\frac{P_j^{vi,e}}{P^e}\right)} \tag{46}$$

And the calculation for *hospitalization*, *ICU admission* and *deaths* under the assumption that the entire population over 12 years of age was fully vaccinated, would be as follows

$$\widehat{H}_j^e = \frac{\widehat{h}_j^{vc,e} P^e}{100.000} \tag{47}$$

$$\widehat{U}_j^e = \frac{\widehat{u}_j^{vc,e} P^e}{100.000} \tag{48}$$

$$\widehat{F}_j^e = \frac{\widehat{f}_j^{vc,e} P^e}{100.000} \tag{49}$$

B Estimation of mean relative incidence and population by age according to vaccination status

Our starting point is Table 10 of *update no. 529 of the Bulletin of December, 23, 2021* of the Center for the Coordination of Health Alerts and Emergencies of the Ministry of Health. From this Table we have prepared Table A1 showing the infections, hospitalizations, ICU admissions and deaths in the eight weeks between October, 18, 2021 and December, 12, 2021. Table A2 also shows the average weekly incidence rates and the relative rates between fully vaccinated and unvaccinated cases

Table A1. Number of cases between October, 18, 2021 and December, 12, 2021. Vaccinated population and non-vaccinated population

Age	$\sum_{j\in\widetilde{s}} Z_j^{vc,e}$	$\sum_{j\in\widetilde{s}} Z_j^{vn,e}$	$\sum_{j\in\widetilde{s}} H_j^{vc,e}$	$\sum_{j\in\widetilde{s}} H_j^{vn,e}$	$\sum_{j\in\widetilde{s}} U_j^{vc,e}$	$\sum_{j\in\widetilde{s}} U_j^{vn,e}$	$\sum_{j\in\widetilde{s}} F_j^{vc,e}$	$\sum_{j\in\widetilde{s}} F_j^{vn,e}$
12 to 29	33,230	18,079	109	197	9	14	2	1
30 to 59	141,660	35,933	1,473	1,532	135	235	45	32
60 to 79	53,775	7,354	3,753	1,199	508	262	271	101
Over 80	9,969	1,069	2,320	390	48	18	511	124

Source: Ministry of Health

Table A2. Average weekly and relative incidence rates by age group

Age	$z^{vc,e}$	$z^{vn,e}$	x_z^e	$h^{vc,e}$	$h^{vn,e}$	x_h^e	$u^{vc,e}$	$u^{vn,e}$	x_u^e	$f^{vc,e}$	$f^{vn,e}$	x_f^e
12 to 29	60	158	0.38	0.2	1.73	0.12	0.02	0.12	0.17	0.00	0.01	0.00
30 to 59	99	177	0.56	1.03	7.57	0.14	0.09	1.16	0.08	0.03	0.16	0.19
60 to 79	75	498	0.15	5.27	81.2	0.06	0.71	17.8	0.04	0.38	6.84	0.06
Over 80	44	222	0.20	10.2	81.1	0.13	0.21	3.75	0.06	2.28	25.8	0.09

Source: Ministry of Health and Own Elaboration. Data from 18/10/2021 to 12/12/2021

From the number of cases and the average weekly incidence, we can estimate the average vaccinated population during the eight-week period as follows

$$\widehat{P}^{vc,e} = \frac{\sum_{j\in\widetilde{s}} Z_j^{vc,e}}{8z^{vc,e}} 100,000 \tag{50}$$

and the non-vaccinated population as

$$\widehat{P}^{vn,e} = \frac{\sum_{j\in\widetilde{s}} Z_j^{vn,e}}{8z^{vn,e}} 100,000 \tag{51}$$

Table A3 provides the estimate of the average proportion of the unvaccinated population in the sum of vaccinated and unvaccinated for the period October, 18, 2021 and December, 12, 2021 consistent with Table 10 of Update No. 529 of the Bulletin of day 23/12/2021.

Table A3. Ratio of unvaccinated to vaccinated and unvaccinated population.. Average for the period from 18/10/2021 to 12/12/2021. Percentage

Age	$\frac{\widehat{P}^{vn,e}}{\widehat{P}^{vc,e} + \widehat{P}^{vn,e}}$
12 to 29	17.1
30 to 59	12.4
60 to 79	2.0
Over 80	2.1

Source: Own Elaboration

The Ministry of Health, in its "daily activity reports on the integrated management of COVID-19 vaccination", offers from April 1, 2021 the data by age group of the stock over time of the total vaccinated population with complete vaccination status $\widetilde{P}_j^{vc,e}$ and population with at least half of the full doses $\widetilde{P}_j^{vc+vi,e}$ (which also includes the population that is fully vaccinated). From these data we can estimate our stock of population that is partially vaccinated.

$$\widetilde{P}_j^{vi,e} = \widetilde{P}_j^{vc+vi,e} - \widetilde{P}_j^{vc,e} \tag{52}$$

and then the unvaccinated population would be obtained by the difference between the total population by age, the fully vaccinated population, and partially vaccinated population.

$$\widetilde{P}_j^{vn,e} = P^e - \widetilde{P}_j^{vc,e} - \widetilde{P}_j^{vi,e} \tag{53}$$

where P^e is the population by age group according to INE data as of January 1, 2020.

Table A4. Population in Spain by age

Age	Population
12 a 29[1]	8,847,760
30 a 59	21,155,446
60 a 79	9,297,031
Over 80	2,834,024

Source: INE and Own Elaboration
1 The INE establishes the cutoff at ages 10 and 15.
It has been assumed that the population aged 12 to 14 years old accounts for 3/5 of the population interval of 10 to 14 years.

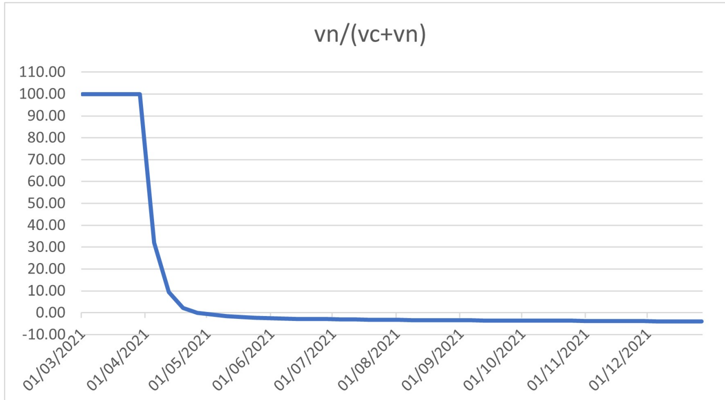


Figure B.1: *Ratio of unvaccinated to the sum of unvaccinated and fully vaccinated (over 80 years old)*

The $\widetilde{P}_j^{vc,e}$, $\widetilde{P}_j^{vn,e}$, and $\widetilde{P}_j^{vi,e}$ series have some problems of consistency both with the population by age groups provided by the INE (Table A.4), and with the ratio of unvaccinated to unvaccinated plus fully vaccinated in Table A3. For example, Figure B.1 represents this ratio for the population over 80 years of age ($\frac{\widetilde{P}^{vn,e}}{\widetilde{P}^{vc,e}+\widetilde{P}^{vn,e}}$)

In view of this graph, the ratio becomes negative because, since May, the number of people who have received at least one dose (which includes those who are fully vaccinated) exceeds the population over 80 years of age according to the INE.

To correct for these level problems, maintaining the time variation rate of the series provided by the Ministry of Health (Ministerio de Sanidad, 2021b) we multiplied $\widetilde{P}_j^{vc,e}$ and $\widetilde{P}_j^{vc+vi,e}$ by a scalar τ such that, after following the procedure described for obtaining $\widehat{P}^{vn,e}$ (let us denote by $P^{vn,e}$ the new series obtained when the procedure is applied by multiplying $\widetilde{P}_j^{vc,e}$ and $\widetilde{P}_j^{vc+vi,e}$ by a scalar τ), the average ratio $\frac{P_j^{vn,e}}{\tau\widehat{P}_j^{vc,e}+P_j^{vn,e}}$ for the period October, 18, 2021 to December, 12, 2021 replicates the values in Table A3. Let us call $P_j^{vc,e}=\tau\widehat{P}_j^{vc,e}$ and $P_j^{vi,e}=\tau\left(\widetilde{P}_j^{vc+vi,e}-\widehat{P}_j^{vc,e}\right)$. Table A5 shows the level correction performed on the series $\widetilde{P}_j^{vc,e}$ and $\widetilde{P}_j^{vc+vi,e}$ through the parameter τ as well as the ratio $\frac{\widetilde{P}^{vn,e}}{\widetilde{P}^{vc,e}+\widetilde{P}^{vn,e}}$ obtained with the original data, and the ratio $\frac{P_j^{vn,e}}{P_j^{vc,e}+P_j^{vn,e}}$ resulting from the corrected series and which is the one estimated in Table A3.//

As can be seen, the main inconsistency occurs in people over 80, while in those younger than 60 the weekly series of the RENAVE (2021) are consistent with those published in Table 10 of the Report on the Vaccination Status of Notified Cases in SiViEs.

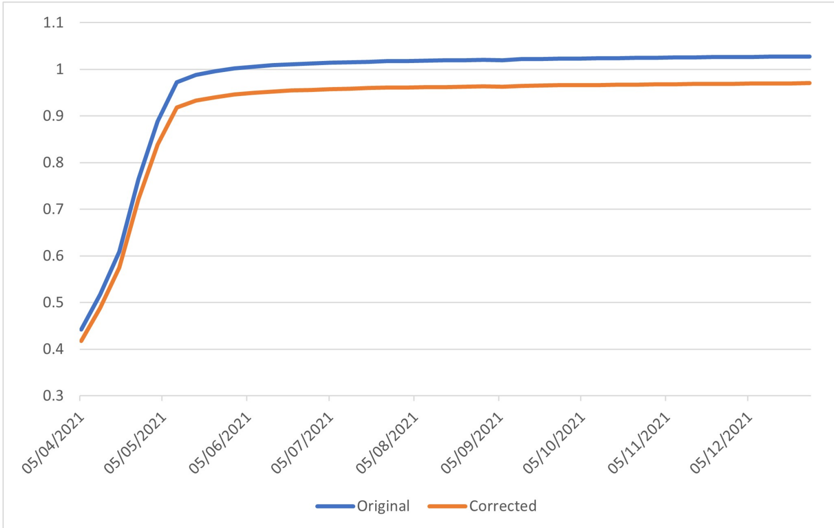


Figure B.2: : *Proportion of the population over 80 years of age that is fully vaccinated.*

Table A5. Series correction factor

Age		τ	$\frac{\widetilde{P}^{vn,e}}{\widetilde{P}^{vc,e}+\widetilde{P}^{vn,e}}$	$\frac{P^{vn,e}}{P^{vc,e}+P^{vn,e}}$
12 to 29		0.9992	17.05	17.12
30 to 59		0.9970	12.15	12.42
60 to 79		0.9866	0.69	2.02
Over 80		0.9442	-3.77	2.08

Source: Own Elaboration

Graph B.2 shows the evolution of the proportion of fully vaccinated persons in the population over 80 years of age, with the original data and with the correction made.

Working paper

2022

22/03 **José E. Boscá, José Cano, Javier Ferri:** Covid-19 in Spain during 2021: what have vaccines achieved and what is the health cost of vaccine hesitancy?

22/02 **Adriana Haring and Mario Iparraguirre:** Argentina. Futuro de los sectores post pandemia.

22/01 **Ángel de la Fuente:** Series largas de algunos agregados económicos y demográficos regionales: actualización de RegData hasta 2020 (RegData y RegData Dem versión 6.1-2020).

2021

21/11 **Ángel de la Fuente and Rafael Doménech:** Cross-country data on skills and the quality of schooling: a selective survey.

21/10 **Ángel de la Fuente:** La evolución de la financiación de las comunidades autónomas de régimen común, 2002-2019.

21/09 **Ángel de la Fuente:** La liquidación de 2019 del sistema de financiación de las comunidades autónomas de régimen común.

21/08 **Rodolfo Méndez-Marcano:** A global vector autoregressive model for banking stress testing.

21/07 **Ali B. Barlas, Seda Guler Mert, Berk Orkun Isa, Alvaro Ortiz, Tomasa Rodrigo, Baris Soybilgen and Ege Yazgan:** Big Data Information and Nowcasting: Consumption and Investment from Bank Transactions in Turkey.

21/06 **Ángel de la Fuente y Rafael Doménech:** El nivel educativo de la población en España y sus regiones: actualización hasta 2019.

21/05 **Saidé Salazar, Jaime Oliver, Álvaro Ortiz, Tomasa Rodrigo and Ignacio Tamarit:**
ESP/ Patrones de Consumo de Efectivo vs Tarjeta en México: una aproximación Big Data.
ING / Cash Vs Card Consumption Patterns in Mexico: A Machine Learning Approach.

21/04 **Ángel de la Fuente:** La financiación autonómica en 2020: una primera aproximación y una propuesta de cara a 2021.

21/03 **Ángel de la Fuente:** Las finanzas autonómicas en 2020 y entre 2003 y 2020.

21/02 **Joxe Mari Barrutiabengoa, J. Julián Cubero and Rodolfo Méndez-Marcano:** Output-side GHG Emissions Intensity: A consistent international indicator.

21/01 **Ángel de la Fuente y Pep Ruiz:** Series largas de VAB y empleo regional por sectores, 1955-2019 Actualización de *RegData-Sect* hasta 2019.

CLICK HERE TO ACCESS THE WORKING DOCUMENTS PUBLISHED IN
Spanish and English

DISCLAIMER

The present document does not constitute an “Investment Recommendation”, as defined in Regulation (EU) No 596/2014 of the European Parliament and of the Council of 16 April 2014 on market abuse (“MAR”). In particular, this document does not constitute “Investment Research” nor “Marketing Material”, for the purposes of article 36 of the Regulation (EU) 2017/565 of 25 April 2016 supplementing Directive 2014/65/EU of the European Parliament and of the Council as regards organisational requirements and operating conditions for investment firms and defined terms for the purposes of that Directive (MIFID II).

Readers should be aware that under no circumstances should they base their investment decisions on the information contained in this document. Those persons or entities offering investment products to these potential investors are legally required to provide the information needed for them to take an appropriate investment decision.

This document has been prepared by BBVA Research Department. It is provided for information purposes only and expresses data or opinions regarding the date of issue of the report, prepared by BBVA or obtained from or based on sources we consider to be reliable, and have not been independently verified by BBVA. Therefore, BBVA offers no warranty, either express or implicit, regarding its accuracy, integrity or correctness.

This document and its contents are subject to changes without prior notice depending on variables such as the economic context or market fluctuations. BBVA is not responsible for updating these contents or for giving notice of such changes.

BBVA accepts no liability for any loss, direct or indirect, that may result from the use of this document or its contents.

This document and its contents do not constitute an offer, invitation or solicitation to purchase, divest or enter into any interest in financial assets or instruments. Neither shall this document nor its contents form the basis of any contract, commitment or decision of any kind.

The content of this document is protected by intellectual property laws. Reproduction, transformation, distribution, public communication, making available, extraction, reuse, forwarding or use of any nature by any means or process is prohibited, except in cases where it is legally permitted or expressly authorised by BBVA on its website www.bbvaresearch.com.

ENQUIRIES TO:
BBVA Research: Azul Street. 4. La Vela Building – 4th and 5th floor. 28050 Madrid (Spain).
Tel.:+34 91 374 60 00 y +34 91 537 70 00 / Fax:+34 91 374 30 25
bbvaresearch@bbva.com www.bbvaresearch.com