# Linear Regression Analysis of Harmonized IgG Antibody Levels Against the SARS-CoV-2 Spike Protein: A Cohort Study in Healthcare Workers

Ana Leonor Saraiva[1,2] (ana.saraiva@insa.min-saude.pt), Vera Afreixo[1,3] (vera@ua.pt), Ausenda Machado[2,4] (ausenda.machado@insa.min-saude.pt) and Vânia Gaio[2,4] (vania.gaio@insa.min-saude.pt)

[1] Department of Mathematics (DMAT), University of Aveiro, Aveiro, Portugal
[2] Department of Epidemiology. Instituto Nacional de Saúde Doutor Ricardo Jorge. Lisbon
[3] Center for Research and Development in Mathematics and Applications (CIDMA), University of Aveiro, Aveiro, Portugal
[4] Public Health Research Center. Escola Nacional de Saúde Pública. Universidade NOVA de Lisboa. Lisbon. & Comprehensive Health Research Center. Universidade NOVA de Lisboa. Lisbon

## INTRODUCTION

The contact of healthcare workers with patients places them at a higher risk of infection. Making their vaccination a global and national priority.

Monitoring vaccine performance is therefore essential. Measuring IgG antibodies against the SARS-CoV-2 Spike protein is a key approach to assessing immunity over time.

The rapid development of independent antibody assays by different manufacturers resulted in non-comparable outcomes.

## METHODOLOGY

In 2021, a prospective cohort study was launched to assess COVID-19 vaccine effectiveness against symptomatic infection among healthcare workers from three hospitals, Centro Hospitalar de Lisboa Ocidental (CHLO), Centro Hospitalar Tondela-Viseu (CHTV), and Centro Hospitalar e Universitário de Coimbra (CHUC).

**Serological data were collected at six time points:**

Pre-vaccination → Post-vaccination → 3 months → 6 months → 12 months → Post-booster

### STEP 1: SORTING THE DATA
The data were sorted to allow the formation of observation pairs between methods.

### STEP 2: NORMALIZATION AND LINEARIZATION
The bestNormalize package was used to normalize and linearize the data by timepoint and hospital center.

### STEP 3: QUANTILE INTERPOLATION
Quantiles were calculated (rank function), and two methods were applied for interpolation: cubic and linear.

### STEP 4: APPLICATION OF REGRESSION
Deming regression was applied, given that it assumes measurement errors in both variables.

### STEP 5: COMPARISON OF HARMONIZED TITRES
Data were transformed using the Deming regressions, then geometric mean was calculated per time point.

### STEP 6: LINEAR REGRESSION ANALYSIS
After harmonization, three linear regressions were fitted

The study aimed to harmonize serological data across these hospitals and to model antibody increases and decreases over time using linear regression.

## RESULTS & DISCUSSION

After harmonization, three linear regressions were fitted: one for the increase between pre-vaccination and post-vaccination, another for the decrease between post-vaccination and 12 months after vaccination, and finally, one for the increase between 12 months after vaccination and after the booster dose.

### LINEAR REGRESSIONS

'PRE-VACCINATION'-'POST-VACCINATION'

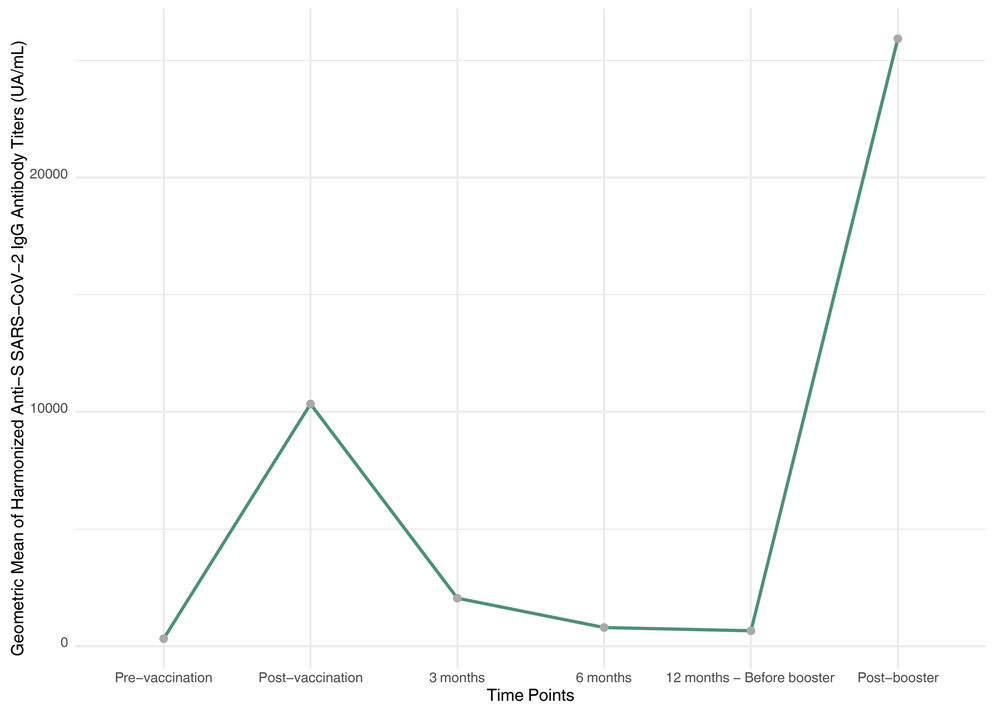
| Characteristics | Coefficient | 95% CI | P-value |
|---|---|---|---|
| **Age** | | | |
| ≤50 | ----- | ----- | ----- |
| >50 | -2795 | -13 862; 8271 | 0.609 |
| **Chronic condition** | | | |
| No chronic condition | ----- | ----- | ----- |
| Chronic condition | 918 | -11 523; 13 359 | 0.881 |
| **Smoking status** | | | |
| Non-smoker | ----- | ----- | ----- |
| Smoker | 9696 | -4480; 23 872 | 0.172 |
| **Contact with COVID patients** | | | |
| No contact | ----- | ----- | ----- |
| Contact | 8441 | -7360; 24 243 | 0.283 |
| **Previous infection** | | | |
| No previous infection | ----- | ----- | ----- |
| Previous infection | -6852 | -17 940; 4235 | 0.216 |

'POST-VACCINATION'-'12 MONTHS'

| Characteristics | Coefficient | 95% CI | P-value |
|---|---|---|---|
| **Hospital** | | | |
| CHUC | ----- | ----- | ----- |
| CHLO | 17 | 3; 30 | **0.015** |
| **Age** | | | |
| ≤50 | ----- | ----- | ----- |
| >50 | -7 | -21; 7 | 0.326 |
| **Chronic condition** | | | |
| No chronic condition | ----- | ----- | ----- |
| Chronic condition | 7 | -8; 21 | 0.387 |
| **Smoking status** | | | |
| Non-smoker | ----- | ----- | ----- |
| Smoker | 5 | -12; 21 | 0.595 |
| **Contact with COVID patients** | | | |
| No contact | ----- | ----- | ----- |
| Contact | 3 | -15; 21 | 0.744 |
| **Previous infection** | | | |
| No previous infection | ----- | ----- | ----- |
| Previous infection | -20 | -62, 22 | 0.359 |

'12 MONTHS'-'POST-BOOSTER'

| Characteristics | Coefficient | 95% CI | P-value |
|---|---|---|---|
| **Hospital** | | | |
| CHUC | ----- | ----- | ----- |
| CHLO | -3 101 778 | -3 307 775; -2 895 781 | **<0.001** |
| **Age** | | | |
| ≤50 | ----- | ----- | ----- |
| >50 | 811 550 | 598 774; 1 024 327 | **<0.001** |
| **Chronic condition** | | | |
| No chronic condition | ----- | ----- | ----- |
| Chronic condition | 129 484 | -93 025; 351 994 | 0.254 |
| **Smoking status** | | | |
| Non-smoker | ----- | ----- | ----- |
| Smoker | 230 939 | -27 369; 489.246 | 0.08 |
| **Contact with COVID patients** | | | |
| No contact | ----- | ----- | ----- |
| Contact | -225 704 | -509 589; 57 981 | 0.119 |
| **Previous infection** | | | |
| No previous infection | ----- | ----- | ----- |
| Previous infection | -307 195 | -667 592; 53 202 | 0.095 |




In the linear regression analysis for "Post-vaccination" – "12 months – Before booster" only the variable "Hospital" showed a statistically significant association with the percentage decrease in IgG antibodies between the two time points analyzed.

Individuals followed at CHLO showed, on average, a 17 percentage point greater decrease compared to CHUC, suggesting a higher antibody decline at this center.

Two variables showed statistical significance in the linear regression "12 months" – "Post-booster": the variable identifying the hospital center and age.

Notably, individuals from CHLO showed, on average, a significantly smaller percentage increase in antibodies after the booster dose compared to those from CHUC.

Despite the overall increase in IgG antibodies after the booster dose, the immune response varies across individuals.

Individuals over 50 years old showed a higher percentage increase after the third dose, possibly because they started from lower baseline levels, unlike younger individuals who already had higher values.

## CONCLUSIONS

The proposed methodology shows promise for harmonizing multicenter data, ensuring comparability and reliability in national and European studies, and enabling a clear evaluation of immune response dynamics among healthcare workers.

Vaccination and booster doses significantly increased antibody levels, while differences between hospitals and individual characteristics influenced the magnitude of these responses. These findings improve our understanding of humoral immunity and may guide future vaccination strategies.

**Funding:** The data of the study were originally collected as part of the project 'Developing an infrastructure and performing vaccine effectiveness studies for COVID- 19 vaccine in the EU/EEA' (Contract ECD.11486 Lot3 (HCW) and amendment Nº ECD.11486), and the Enhanced laboratory support to perform assessment of vaccine effectiveness against SARS- CoV- 2 infection (ECD.12175) and the 'Vaccine Effectiveness, Burden and Impact Studies (VEBIS) of COVID- 19 and Influenza', funded by the European Centre for Disease Prevention and Control through a service contract with Epiconcept (ECD.12609).