International Journal of Advanced Trends in Engineering and Technology (IJATET)
International Peer Reviewed - Refereed Research Journal, Website: www.dvpublication.com
Impact Factor: 5.965, ISSN (Online): 2456 - 4664, Volume 11, Issue 2, July - December, 2026

PUBLIC HEALTH RESPONSES AND REPORTED COVID-19 INCIDENCE: TESTING INTENSITY, VACCINATION COVERAGE, AND GOVERNMENT POLICY STRINGENCY

Divya Vishwanath* & Paul Johnson Asamoah**
* Janardhan Rai Nagar Rajasthan Vidyapeeth, Udaipur, Rajasthan, India
** Power Life University College, Accra, Ghana, West Africa

Cite This Article: Divya Vishwanath & Paul Johnson Asamoah, "Public Health Responses and Reported Covid-19 Incidence: Testing Intensity, Vaccination Coverage, and Government Policy Stringency", International Journal of Advanced Trends in Engineering and Technology, Volume 11, Issue 2, July - December, Page Number 23-32, 2026.
Copy Right: © DV Publication, 2026 (All Rights Reserved). This is an Open Access Article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited.
Type of Review: Peer Reviewed as per |C|O|P|E| Guidance.
Disclaimer: The scholarly papers reviewed and published by DV Publication, India, reflect the views and opinions of their respective authors and do not necessarily represent the views or opinions of DV Publication. The publisher disclaims any responsibility for any harm, loss, or damage resulting from the use of the published content by any party.
DOI: https://doi.org/10.5281/zenodo.24041164

Abstract:

This analysis examines how public health responses relate to reported COVID 19 incidence, addressing the distinction between disease burden and surveillance visibility in international comparisons. An unbalanced panel of 756 country month observations across 76 countries from December 2020 to June 2022 was analysed using descriptive statistics, pooled Spearman correlations and 4,000 country cluster bootstrap replications. Testing intensity showed the strongest positive association with reported incidence, with a correlation of 0.571, followed by vaccination coverage at 0.278; government policy stringency showed a negative association of −0.133. All three bootstrap confidence intervals excluded zero. Association directions persisted when observations were restricted to consistent testing definitions and more complete monthly coverage. These unadjusted relationships cannot distinguish increased case detection from epidemic pressure or identify intervention effects. The contribution is a measurement aligned comparison demonstrating that diagnostic activity, preventive uptake and formal restrictions have distinct observational relationships with recorded burden. The findings sharpen the distinction between intervention effectiveness and surveillance outcomes, supporting international benchmarking that explicitly accounts for reporting definitions and observation completeness.

Key Words: Disease Surveillance; Testing Intensity; Vaccination Coverage; Government Stringency; Cluster Bootstrap.

1. Introduction:

Reported COVID 19 incidence reflects both the spread of infection and the capacity of health systems to detect it. This distinction matters for international comparisons because countries can record different case rates despite similar underlying infection burdens. Expanded testing can increase recorded incidence by identifying previously undetected infections, while limited diagnostic access can suppress the apparent burden. Hasell et al., 2020, document differences in national testing definitions and reporting practices that complicate comparisons of confirmed cases. Evaluating public health responses against reported incidence therefore requires attention to how the outcome is produced, not only to the interventions intended to reduce transmission. Behavioural epidemiology provides a theoretical basis for examining these relationships. Epidemic conditions influence individual decisions and institutional responses, which can subsequently alter exposure and transmission. Funk, Salathé, and Jansen, 2010, explain how behavioural adaptation creates feedback between disease dynamics, perceived risk and preventive action. Testing intensity, vaccination coverage and government policy stringency occupy different positions within this system. Testing influences case ascertainment and can support isolation; vaccination can reduce susceptibility; restrictions can constrain opportunities for infectious contact. Rising infections can also increase testing demand and prompt stronger restrictions. Consequently, contemporaneous associations need not reproduce the direction of an intervention's preventive effect.

The testing literature illustrates this ambiguity. Larremore et al., 2021, show through modelling that screening frequency and reporting speed influence the capacity of testing programmes to interrupt transmission. However, a population adjusted testing count does not measure diagnostic turnaround time, isolation adherence or the targeting of screening. Higher testing intensity can therefore coincide with higher reported incidence through broader detection, greater epidemic pressure or both. An international comparison based on observed testing volumes can characterise this relationship, but it cannot independently identify the contribution of each mechanism. Vaccination introduces a related distinction between intervention effectiveness and aggregate coverage. Dagan et al., 2021, provide evidence from Israel that vaccination reduced documented infection in a nationwide vaccination setting. Feikin et al., 2022, demonstrate through systematic review and meta regression that protection against infection changes with time since vaccination. These findings establish why initial protocol completion cannot be treated as a direct measure of effective protection at every observation date. Countries can experience increasing coverage alongside epidemic waves, and national averages do not reproduce comparisons between otherwise comparable vaccinated and unvaccinated individuals. A positive aggregate association therefore requires interpretation at the population and temporal scale on which it was calculated.

Government restrictions operate through institutional constraints on contact opportunities, but formal policy intensity does not directly measure implementation. Haug et al., 2020, show that intervention effectiveness varies across policy combinations and contexts. Their analysis concerns interventions and transmission dynamics, whereas a composite stringency index summarises recorded restrictions. The distinction matters because similar index values can represent different policy packages, and governments may strengthen restrictions in response to rising cases. Comparing stringency with contemporaneous reported incidence consequently addresses an observational relationship rather than the number of infections prevented by policy.

The reviewed evidence establishes important preventive mechanisms, but it does not determine the pooled relationships among these three observed response measures and reported incidence in the analytical sample examined here. The contribution is therefore bounded: this analysis places diagnostic activity, initial vaccination coverage and formal restrictions alongside the same surveillance outcome, using identical eligible days within each country month. It then examines whether the observed associations persist when testing definitions are more consistent and monthly observation coverage is more complete. This approach addresses comparability within the assembled data without claiming to resolve differences in national surveillance systems or identify causal intervention effects.

This study aims to characterise the associations between testing intensity, vaccination coverage, government policy stringency and reported COVID 19 incidence across countries. The central research question asks how each response measure relates to recorded incidence and whether those relationships change under alternative measurement and coverage restrictions. The specific objectives are to describe the distributions of the four measures, estimate their pooled rank associations and assess sensitivity to testing definitions and monthly completeness. No moderator, mediator or combined response index is constructed. The empirical analysis uses 756 country month observations across 76 countries between December 2020 and June 2022. Monthly measures average the same jointly observed days, with at least 15 eligible days required for inclusion. Descriptive statistics and exploratory Spearman correlations characterise the observed relationships, while country cluster bootstrap intervals account for repeated observations within countries. Sensitivity comparisons restrict the sample to tests performed and, separately, to months with at least 25 jointly observed days. Testing and vaccination retain positive associations with reported incidence across these comparisons, while stringency retains a negative association. These findings contribute a transparent international assessment of surveillance relationships and their measurement boundaries. Their practical relevance lies in informing comparisons of recorded epidemic burden alongside diagnostic activity and reporting coverage, rather than ranking national intervention effectiveness.

2. Theory, Conceptual Framework and Research Questions Development:

This analysis draws on behavioural epidemiology to explain reported COVID 19 incidence as the joint outcome of infection processes, institutional responses and case detection. Individuals adjust contact behaviour as perceived infection risk changes, while governments respond through testing, vaccination and restrictions. These responses can influence transmission, but rising infections can also intensify intervention. Funk, Salathé, and Jansen, 2010, explain how behavioural responses create feedback between epidemic conditions and individual decisions. This reciprocal relationship makes contemporaneous associations between public health responses and reported incidence theoretically ambiguous. The framework distinguishes disease occurrence from disease observation. Reported incidence measures detected and recorded cases rather than all infections. Testing affects the probability of observation, vaccination influences susceptibility, and policy stringency represents formal restrictions on opportunities for contact. These measures therefore capture distinct processes rather than interchangeable indicators of a single latent construct. The available panel contains 756 country month observations across 76 countries between December 2020 and June 2022. Each observation averages the four measures over the same eligible days, with at least 15 jointly observed days per month.

Testing intensity measures diagnostic activity relative to population size. Its immediate connection to reported incidence operates through ascertainment: increasing diagnostic coverage can reveal infections that would otherwise remain unrecorded. However, testing volumes also respond to symptomatic demand and outbreak investigation. Hasell et al., 2020, document differences in national testing definitions and reporting practices that constrain comparisons across countries. Consequently, higher recorded testing intensity can indicate broader surveillance, greater epidemic pressure, or both. Testing can also contribute to transmission reduction when timely diagnosis leads to isolation. Larremore et al., 2021, demonstrate through modelling that screening frequency and reporting speed strongly influence the capacity of testing programmes to interrupt transmission. This mechanism differs from ascertainment because it can reduce subsequent infections while initially increasing detected cases. The workbook measures testing volume but does not measure diagnostic turnaround time, isolation adherence or screening design. These mechanisms therefore explain competing theoretical expectations without becoming additional empirical variables. Because detection and prevention can operate in opposite directions, this analysis does not impose a universal positive or negative relationship between testing and reported incidence. It evaluates the observed association while retaining the distinction between increased detection and increased transmission.

RQ1: How is testing intensity associated with reported COVID 19 incidence across the observed country months?

Vaccination coverage represents the population share that completed the initial vaccination protocol. Unlike testing intensity, which directly influences case ascertainment, vaccination primarily acts through biological protection. Mathieu et al., 2021, establish a comparable international measurement framework for vaccination uptake, including completion of the initial protocol. Nevertheless, this coverage measure records vaccination status rather than the level of protection remaining at a particular time. The theoretical expectation is that greater effective protection reduces susceptibility and therefore limits infections that can become reported cases. Dagan et al., 2021, provide evidence that vaccination reduced documented infection in a nationwide vaccination setting. That finding supports the protective mechanism, but its magnitude cannot be transferred directly to comparisons of national monthly averages. Individual vaccination effectiveness and the association between aggregate coverage and national incidence concern different units of analysis. The available coverage measure does not distinguish vaccine products, booster uptake, time since vaccination or circulating variants. These factors can change the relationship between recorded coverage and effective protection. Countries can also experience epidemic waves while coverage increases. The analysis therefore examines the observed relationship without interpreting a weak or positive aggregate association as evidence that vaccination increases individual infection risk.

RQ2: How is completion of the initial vaccination protocol associated with reported COVID 19 incidence across the observed country months?

Government policy stringency captures formal restrictions intended to reduce opportunities for infectious contact. Restrictions on workplaces, schools, gatherings and movement can alter contact networks and constrain transmission. Hale et al., 2021, describe the policy database underlying the stringency index and its systematic measurement of government responses. The

index measures announced policy intensity rather than actual compliance, enforcement or reductions in interpersonal contact. The expected preventive mechanism depends on implementation and behavioural response. Haug et al., 2020, show that intervention effectiveness varies with policy combinations and national context. A higher composite index therefore does not necessarily imply a proportionately greater reduction in transmission. Different combinations of restrictions can produce similar index values while imposing different constraints on behaviour. Policy responsiveness introduces a competing relationship. Governments can strengthen restrictions when cases rise and relax them when conditions improve. Higher stringency can therefore coincide with higher reported incidence even when restrictions subsequently reduce transmission. Because the workbook contains monthly averages rather than a design identifying policy effects, the framework treats stringency as an observed institutional response whose association with incidence remains open to empirical assessment.

RQ3: How is government policy stringency associated with reported COVID 19 incidence across the observed country months?

The conceptual framework links three observed explanatory measures, testing intensity, vaccination coverage and government policy stringency, to one observed outcome, reported COVID 19 incidence. Testing captures diagnostic activity, vaccination captures initial protocol coverage, and stringency captures formal restrictions. Their relationships with incidence represent observational associations rather than identified causal effects. No mediator, moderator or latent response index is constructed. The research questions allow positive, negative, negligible or inconclusive associations and do not presume that the available data establish intervention effectiveness.

3. Data:

This section establishes data provenance, documents sample construction and defines the four observed measures used in the analysis. It reports verified coverage, missingness and descriptive statistics while distinguishing completed data preparation from statistical procedures that have not been implemented.

3.1 Data Sources, Population, and Sample Construction:

The analysis uses the archived Our World in Data COVID 19 dataset, whose repository records its last update as 19 August 2024. The source integrates confirmed cases from the World Health Organization, testing and vaccination records compiled by Our World in Data, and government policy indicators from the Oxford COVID 19 Government Response Tracker. The archived file is identified by repository commit f0d53320b19703cf565940d8269181a8423d7032. The extraction timestamp and source file checksum are recorded in the accompanying workbook. Our World in Data, 2024, provides the source documentation and access route. Testing data follow the international compilation described by Hasell et al., 2020. Vaccination data follow the reporting framework documented by Mathieu et al., 2021. Policy stringency originates from the database described by Hale et al., 2021. These sources provide population adjusted measures, but national reporting definitions and administrative practices remain relevant to comparability. Our World in Data authored material is distributed under Creative Commons Attribution terms; data supplied by other organisations retain their original restrictions.

The target population comprises national public health systems reporting COVID 19 surveillance and response indicators. The accessible frame comprises country and territory records in the archived source, excluding identifiers beginning with OWID_. The geographical setting is international, and the sectoral setting is public health surveillance. Selection follows data availability and explicit eligibility rules rather than probability sampling. The retained sample therefore represents reporting systems with sufficient overlapping information, rather than a representative sample of all national health systems. The source contains daily observations, while the analytical unit is the country month. The final dataset forms an unbalanced panel covering December 2020 to June 2022, with eligible daily observations extending from 13 December 2020 to 23 June 2022. It does not provide continuous empirical coverage of the broader 2010 to 2025 scope. Testing updates ended on 23 June 2022, as documented by Our World in Data, 2022.

Table 1: Verified sample construction

Stage	Geographical Identifiers Retained	Observations Retained	Implemented Rule
Accessible source frame	237	395,311 daily records	Country and territory identifiers within the requested period, excluding OWID_ aggregates
Complete observations	117	23,577 daily records	Remove 371,734 records missing at least one required measure
Valid observations	117	23,556 daily records	Remove 21 records with vaccination coverage above 100 percent
Monthly aggregation	117	1,386 country months	Group valid daily records by country identifier and calendar month
Minimum coverage restriction	76	756 country months	Exclude 630 groups containing fewer than 15 jointly observed days
Final descriptive sample	76	756 country months	Retain all four measures without additional model restrictions

The final sample incorporates 20,401 eligible daily observations. Countries contribute between one and 19 monthly observations. Monthly geographical coverage ranges from one country in December 2020 to 60 countries in June 2021. These differences reflect overlapping reporting availability and prevent interpretation of changes in pooled monthly averages as changes within a constant set of countries.

3.2 Data Integration, Cleaning, and Quality Assurance:

All four indicators were extracted from one consolidated source file. No additional external merge was performed. Daily eligibility was evaluated using the country identifier and observation date, followed by aggregation using the country identifier and calendar month. The procedure retained a daily record only when all four measures were observed, all values were

nonnegative, and vaccination coverage and policy stringency did not exceed 100. An identifier and date audit identified 1,408 excess source records associated with repeated keys for East Timor and the Faroe Islands. None passed the joint completeness and validity requirements. The eligible daily data therefore contain no repeated country and date keys, and the final panel contains no duplicate country and month keys. No source priority rule or averaging of conflicting duplicate records was required for the retained sample. For each country month, the procedure summed each measure over jointly eligible days and divided by the number of those days. All four monthly values consequently refer to the same daily support. Months with fewer than 15 eligible days were removed. The resulting measures are observed day averages rather than complete calendar month averages. No additional interpolation, forward filling, zero substitution or statistical imputation was performed. Table 2 reports daily missingness within the 76 retained countries over the actual observation window, before complete observation selection. The year labels represent only the portions of those years within that window.

Table 2: Daily missingness within retained countries

Period	Available Daily Records	Missing Testing Intensity	Missing Vaccination Coverage	Missing Stringency	Missing Reported Incidence
December 2020 observation window	1,444	228	1,380	0	0
January to December 2021	27,740	4,403	8,769	0	0
January to June 2022 observation window	13,224	4,715	4,278	0	0
Total	42,408	9,346	14,427	0	0

Missing counts overlap because a daily record can lack more than one measure. Across retained countries, unavailable testing observations range from zero to 505 days, while unavailable vaccination observations range from zero to 493 days. Missing vaccination records were not interpreted as evidence of zero vaccination. Complete observation selection avoids constructing unsupported values, but it does not establish that missingness is random. Hughes et al., 2019, explain why the consequences of deletion depend on the missingness process and the intended analysis. No trimming or winsorisation was applied to valid extreme observations. Testing intensity has no imposed upper bound because repeated testing can occur. The retained monthly records identify tests performed in 478 observations, people tested in 223, samples tested in 47, and unclear units in eight. Population adjustment does not eliminate these definitional differences. The archived file contains revised historical information rather than a reconstruction of information available at each original reporting date. Source smoothing and revisions were preserved. No prospective forecasting claim is made. The workbook records the implemented selection rules, but it does not contain a dated preregistration or a comparison of alternative coverage thresholds. The audit did not alter the retained sample.

3.3 Variable Measurement, Construction, and Validation:

The framework contains three observed explanatory measures and one observed outcome. Public health responses provide the organising concept, but no combined response index is constructed. Testing intensity measures diagnostic activity, vaccination coverage measures completion of the initial protocol, and stringency measures formal policy restrictions. Reported incidence measures detected cases rather than all infections.

Table 3: Variable dictionary

Variable	Analytical Role	Definition and Construction	Source Identifier	Unit and Frequency	Direction and Missing Rule	Coverage and Limitation
Testing intensity	Explanatory measure	Mean daily new tests relative to population over jointly eligible days	New tests per thousand	Tests per 1,000 people per day, averaged monthly	Higher values indicate greater recorded testing activity; exclude days with missing values	756 observations; national testing units differ
Vaccination coverage	Explanatory measure	Mean population share completing the initial vaccination protocol over jointly eligible days	People fully vaccinated per hundred	Percent of population, averaged monthly	Higher values indicate greater recorded coverage; exclude missing values and values above 100	756 observations; excludes distinctions in boosters, products and time since vaccination
Government policy stringency	Explanatory measure	Mean source composite policy index over jointly eligible days	Stringency index	Index from zero to 100, averaged monthly	Higher values indicate stricter recorded restrictions; exclude missing or invalid values	756 observations; measures formal policy rather than compliance
Reported incidence	Outcome	Mean source seven day smoothed new confirmed cases relative to population over jointly eligible days	New cases smoothed per million	Cases per million people per day, averaged monthly	Higher values indicate greater recorded incidence; exclude missing or negative values	756 observations; depends on detection and reporting practices

Each monthly measure was calculated using Equation 1.

$$\bar{x}_{cm} = \frac{\sum_{d \in D_{cm}} x_{cd}}{n_{cm}}$$

Here, $\bar{x}_{cm}$ denotes the monthly value of a measured indicator for country c and month m. The term x_{cd} denotes its source value on day d. The set D_{cm} contains days meeting all four measurement requirements, and n_{cm} denotes the number of those days. A monthly observation is retained when $n_{cm} \geq 15$. Daily observations receive equal weight within each retained month. Testing and reported incidence originate as daily flow measures. Vaccination coverage originates as a cumulative population share, while stringency records policy conditions at a point in time. Monthly aggregation converts each into a period average. Vaccination coverage is therefore neither an end of month value nor a monthly vaccination flow. Differences in coverage represent percentage points rather than percentage growth.

Table 4: Descriptive statistics for retained country months

Measure	Valid Observations	Mean	Standard Deviation	Median	Minimum	Maximum
Testing per 1,000 people per day	756	4.298	7.271	2.327	0.067	69.962
Fully vaccinated population percent	756	39.523	29.740	40.441	0.000	99.375
Policy stringency index	756	49.421	18.175	49.045	5.560	89.162
Reported cases per million people per day	756	369.949	657.009	147.428	0.369	5,572.739

These statistics assign equal weight to each retained country month. They are not global population weighted estimates. Testing intensity and reported incidence have means above their medians, indicating that larger observations influence pooled averages. Validation established completeness, admissible ranges, unique analytical keys and consistency between monthly sums, eligible day counts and calculated averages. No independent external benchmark comparison was performed.

3.4 Analytical Framework, Identification, and Reproducibility:

The completed analysis is descriptive. Its empirical quantities are the distributions, coverage and variation of the four measures across observed country months. Means and medians describe typical recorded levels, while standard deviations and ranges describe dispersion. No regression coefficients, causal effects, predictive accuracy measures or hypothesis decisions have been estimated. The data permit examination of associations, but contemporaneous measurement does not identify intervention effects. Rising infections can increase testing demand and trigger stronger restrictions. Testing also changes case ascertainment, while vaccination coverage does not measure the protection remaining against each circulating variant. These relationships prevent a direct interpretation of differences in reported incidence as differences caused by public health responses. Rohrer, 2018, distinguishes observed relationships from causal effects requiring additional identifying assumptions. No controls, interactions, fixed effects, lagged variables or inferential procedures were implemented. There are consequently no additional sample losses from estimation requirements. The analysis also contains no alternative threshold assessment or estimated robustness comparison. Unequal country coverage, incomplete monthly support and differences in testing definitions remain limits on international generalisation. The workbook provides the analytical records, monthly source sums, eligible day counts, calculation formulas, measurement definitions and source references. Its checksum identifies the source file used. These materials allow verification of the monthly calculations, although the workbook does not embed the full daily source file or a complete executable extraction archive. Consistent daily eligibility aligns the four measures within each retained country month; it does not resolve differences in national reporting systems or recover observations excluded through missingness.

4. Methods:

4.1 Research Design and Analytical Scope:

This analysis uses an observational design to examine pooled associations between public health responses and reported COVID 19 incidence. The unit of analysis is the country month. The three explanatory measures are testing intensity, vaccination coverage and government policy stringency, as defined in the variable dictionary in Table 3. Each is examined separately against reported incidence. The estimand is the Spearman rank correlation between an explanatory measure and reported incidence across eligible country months. The estimates combine differences between countries and changes across months. They do not isolate variation within countries or adjust for other explanatory measures. No treatment assignment, comparison group or exogenous source of variation supports causal identification. The design addresses whether higher recorded response levels coincide with higher or lower recorded incidence. Interpretation accounts for competing explanations: outbreaks can increase testing demand and policy restrictions, while diagnostic activity can increase case detection. Changes in epidemic conditions can also coincide with vaccination progress. Consequently, the estimates describe unadjusted associations rather than intervention effects. The sample construction documented in Table 1 yields 756 observations across 76 countries between December 2020 and June 2022. The panel is unbalanced, and eligibility depends on overlapping reporting availability. Countries with more retained months contribute more observations to the pooled estimates. Neither the sample size nor its geographical breadth establishes population representativeness.

4.2 Measurement Alignment and Association Estimation:

Monthly measures follow the jointly observed day aggregation defined in Equation 1. This construction aligns all four indicators within each country month. It does not convert incomplete observation coverage into complete calendar month coverage. No further logarithmic transformation, standardisation, lag construction or composite aggregation was applied before association estimation. Descriptive analysis reports valid counts, means, sample standard deviations, medians, quartiles and observed ranges. Standard deviations use the number of observations less one as the denominator. Quartiles use linear interpolation between ordered observations. These summaries characterise dispersion and identify differences between typical values and averages influenced by larger observations. Spearman correlation was selected to describe monotonic relationships without requiring a linear association between the original measurement scales. Its rank transformation reduces sensitivity to extreme magnitudes, although it does not eliminate measurement error or influential observations. De Winter et al., 2016, compare the behaviour of Pearson and Spearman coefficients across distributions and sample sizes, including their sensitivity to outliers. For each explanatory measure, the implemented estimator is expressed in Equation 2.

$$\hat{\rho}_{s,j} = \frac{\sum_{i=1}^{N}\left(R_{ij} - \bar{R}_j\right)\left(S_i - \bar{S}\right)}{\sqrt{\sum_{i=1}^{N}\left(R_{ij} - \bar{R}_j\right)^2 \sum_{i=1}^{N}(S_i - \bar{S})^2}}$$

Here, $\hat{\rho}_{s,j}$ denotes the estimated Spearman correlation for explanatory measure j. The index iidentifies a country month, and Nis the number of observations in the relevant analytical sample. The term R_{ij} denotes the rank of explanatory measure j, while S_i denotes the rank of reported incidence. The terms $\bar{R}_j$ and $\bar{S}$are their respective mean ranks. Tied observations receive average ranks. Equation 2 was estimated separately for each explanatory measure. No controls, fixed effects, interactions or

mediating variables were included. The coefficients therefore cannot be interpreted as independent contributions, marginal effects or shares of an overall response effect. Each retained country month receives equal weight, so the estimates are not population weighted international associations.

4.3 Country Cluster Bootstrap and Computational Implementation:

Uncertainty was estimated through a country cluster bootstrap. Resampling individual country months would separate observations belonging to the same national reporting system. Instead, each replication sampled countries with replacement and retained every eligible monthly observation belonging to each selected country. This preserves the observed temporal grouping within countries. Cameron and Miller, 2015, explain the importance of accounting for dependence within clusters when conducting statistical inference. For a sample containing G countries, each replication drew G country blocks with replacement. A country selected more than once contributed its complete observed block each time. Because country blocks differ in length, the number of country month observations could vary between replications. Ranks were recalculated within every resampled dataset, and the same country draw was used for all three correlations. Each analytical sample used 4,000 replications. The 95 percent percentile interval is defined in Equation 3.

$$\mathrm{CI}_{j,95} = \left[Q_{0.025} \left\{ \hat{\rho}^{*}_{s,j,b} \right\}_{b=1}^{4000} , Q_{0.975} \left\{ \hat{\rho}^{*}_{s,j,b} \right\}_{b=1}^{4000} \right]$$

Here, $\mathrm{CI}_{j,95}$ denotes the interval for explanatory measure j. The term $\hat{\rho}^{*}_{s,j,b}$ is the correlation calculated in bootstrap replication b, and $Q_{0.025}$ and $Q_{0.975}$ denote the corresponding empirical quantiles. Quantiles use linear interpolation. The reported intervals are percentile intervals rather than normal approximation intervals or bias corrected intervals. Calculations used Python 3.12.14, NumPy 2.3.5 and SciPy 1.17.0. Correlations were computed using scipy.stats.spearmanr. Bootstrap sampling used numpy.random.default_rng with seed 73021. Country identifiers were ordered using numpy.unique. One random number generator was initialised and advanced sequentially through the full sample, the testing definition restriction and the monthly coverage restriction. The bootstrap permits dependence among observations within a country but relies on countries as the resampling units. It does not address arbitrary dependence between countries, including common epidemic shocks. Its intervals describe uncertainty under that resampling structure and do not incorporate uncertainty from source revisions, measurement error or the original selection of reporting countries. No additional small cluster correction, multiple comparison adjustment or bootstrap significance test was applied.

4.4 Quality Checks, Sensitivity Analysis and Interpretive Limits:

The analytical sequence began with the cleaned workbook sample. Checks verified that the four measures were numeric and complete and that country and month keys were unique. Monthly calculations were checked against their source sums and eligible day denominators. Valid extreme observations were retained without trimming or winsorisation. No additional imputation, interpolation or replacement of missing observations with zero was undertaken. Two exploratory restrictions examined specific comparability concerns. The first retained only country months classified as tests performed, excluding records described as people tested, samples tested or units unclear. This restriction produced 478 observations across 49 countries. The second retained months with at least 25 jointly observed days, producing 550 observations across 62 countries. Each restriction was applied separately to the full sample, followed by fresh calculation of ranks, correlations and bootstrap intervals. The first comparison addresses variation in testing definitions, while the second examines dependence on monthly observation coverage. Both restrictions also change country and period composition. Differences between their estimates and the full sample therefore combine changes in measurement support with changes in the analytical sample. No common sample decomposition was performed.

Sensitivity assessment compared coefficient direction, magnitude and interval width rather than relying solely on whether intervals excluded zero. The resulting estimates and comparisons appear in Tables 6 and 7. These checks were exploratory and were not documented in a preregistered protocol. They assess persistence under two explicit restrictions without establishing that selection bias, confounding or reporting differences have been removed. The accompanying workbook documents measurement definitions, monthly source sums, eligible day counts, calculation formulas and source references. It does not contain a complete executable archive of the extraction and bootstrap procedures. The implementation details reported here specify the estimator, resampling sequence and computational settings required to reproduce the association calculations from the analytical records. The methodological contribution is transparent application of established descriptive and resampling procedures, rather than a new estimator or causal identification strategy.

5. Findings:

This analysis examines how testing intensity, vaccination coverage and government policy stringency relate to reported COVID 19 incidence. The findings combine distributional evidence with exploratory associations and sensitivity comparisons. Interpretation distinguishes observed relationships from intervention effects and accounts for incomplete coverage and repeated country observations.

5.1 Sample Composition and Distributional Characteristics:

The analytical sample contains 756 country month observations across 76 countries between December 2020 and June 2022. All four measures are observed for every retained record. Each country month averages between 15 and 31 jointly observed days, with an average of 26.99 days. The panel incorporates 20,401 eligible daily observations. Countries contribute between one and 19 monthly records, making the panel unbalanced. Coverage is concentrated in 2021, which contributes 548 observations, compared with 207 in 2022 and one in December 2020. Consequently, pooled statistics describe the observed country months rather than a geographically constant population followed throughout the period. Countries with longer reporting histories contribute more observations.

Table 5: Distribution of observed public health responses and reported incidence

Measure	Observations	Mean	Standard Deviation	Median	First Quartile	Third Quartile	Minimum	Maximum
Testing per 1,000 people per day	756	4.298	7.271	2.327	0.825	4.705	0.067	69.962
Fully vaccinated population percent	756	39.523	29.740	40.441	7.916	66.470	0.000	99.375
Policy stringency index	756	49.421	18.175	49.045	38.351	62.332	5.560	89.162
Reported cases per million people per day	756	369.949	657.009	147.428	47.390	374.292	0.369	5,572.739

Testing intensity displays substantial dispersion. Its mean of 4.298 exceeds its median of 2.327, while the central half of observations lies between 0.825 and 4.705 tests per 1,000 people per day. The maximum of 69.962 indicates that a small number of observations extend far above typical recorded testing levels. These differences make the pooled mean an incomplete description of diagnostic activity. Reported incidence exhibits a similar concentration of larger values. The mean of 369.949 cases per million exceeds the median of 147.428, and the standard deviation reaches 657.009. The central half of observations ranges from 47.390 to 374.292 cases per million per day. This dispersion supports examining rank associations alongside descriptive levels, because rank correlations reduce the influence of extreme magnitudes. Vaccination coverage spans almost the entire percentage scale. Its central half ranges from 7.916 to 66.470 percent, reflecting substantial variation in recorded initial protocol completion across countries and months. Policy stringency is more concentrated around its midpoint, with similar mean and median values. These distributions establish variation in all three explanatory measures without implying that their variation is independent of epidemic conditions.

5.2 Associations with Reported Incidence:

The three research questions were examined using exploratory pooled Spearman rank correlations. Each coefficient measures whether higher values of an explanatory variable coincide with higher or lower reported incidence across observed country months. The calculations do not adjust for country characteristics, calendar month or other explanatory measures. Uncertainty was estimated through 4,000 country cluster bootstrap replications using random seed 73021. Each replication sampled countries with replacement and retained all available monthly observations for each sampled country. The reported intervals are the 2.5th and 97.5th percentiles of the resulting coefficient distributions. This procedure preserves repeated observations within countries but does not adjust for dependence between countries arising from shared epidemic shocks.

Table 6: Exploratory associations with reported COVID 19 incidence

Research Question	Explanatory Measure	Spearman Correlation	Lower 95 Percent Confidence Limit	Upper 95 Percent Confidence Limit
RQ1	Testing intensity	0.571	0.483	0.647
RQ2	Vaccination coverage	0.278	0.180	0.373
RQ3	Government policy stringency	−0.133	−0.231	−0.031

Testing intensity has the largest absolute association with reported incidence. The coefficient of 0.571 indicates that country months with higher recorded testing activity tend to occupy higher positions in the reported incidence distribution. Its confidence interval remains positive. RQ1 therefore identifies a positive pooled association between testing and recorded cases. The result does not distinguish broader detection from testing demand generated by infection waves, and it does not establish that testing increases transmission. Vaccination coverage also has a positive pooled association with reported incidence, with a coefficient of 0.278 and a confidence interval from 0.180 to 0.373. RQ2 therefore identifies a positive relationship within the observed panel. This result concerns aggregate coverage and recorded incidence across countries and months. It does not measure individual vaccine effectiveness or demonstrate that vaccination increases infection risk. The available calculation does not separate vaccination progress from changing epidemic conditions, diagnostic activity or differences between national reporting systems.

Government policy stringency has a negative pooled association with reported incidence. The coefficient of −0.133 indicates a weaker rank relationship than those observed for testing or vaccination. Its confidence interval extends from −0.231 to −0.031. RQ3 therefore identifies a negative association in the full sample. The magnitude remains modest, and the contemporaneous comparison cannot distinguish preventive effects from governments strengthening or relaxing restrictions in response to epidemic conditions. All three intervals exclude zero under the resampling procedure. This statistical separation does not establish causal importance. Spearman coefficients describe rank relationships and cannot be translated directly into additional cases, infections prevented or percentage changes in incidence.

5.3 Sensitivity to Testing Definitions and Monthly Coverage:

Two exploratory restrictions assessed whether the association patterns depended on measurement definitions or incomplete monthly support. The first retained only observations explicitly classified as tests performed, producing 478 country months across 49 countries. The second required at least 25 jointly observed days per month, producing 550 country months across 62 countries. Both comparisons used the same correlation and bootstrap procedures as the full sample.

Table 7: Sensitivity of associations to analytical sample restrictions

Restricted Sample	Measure	Spearman Correlation	Lower 95 Percent Confidence Limit	Upper 95 Percent Confidence Limit
Tests performed only	Testing intensity	0.503	0.389	0.612

Restricted Sample	Measure	Spearman Correlation	Lower 95 Percent Confidence Limit	Upper 95 Percent Confidence Limit
Tests performed only	Vaccination coverage	0.252	0.125	0.377
Tests performed only	Policy stringency	−0.142	−0.259	−0.013
At least 25 jointly observed days	Testing intensity	0.564	0.472	0.646
At least 25 jointly observed days	Vaccination coverage	0.304	0.182	0.416
At least 25 jointly observed days	Policy stringency	−0.203	−0.314	−0.082

Restricting testing records to tests performed reduces the testing correlation from 0.571 to 0.503, while preserving its positive direction. The vaccination coefficient changes from 0.278 to 0.252, and the stringency coefficient changes from −0.133 to −0.142. These results indicate that the observed directions persist when testing definitions are more consistent. They do not eliminate differences in national eligibility rules, diagnostic access or reporting practices. Requiring at least 25 jointly observed days leaves the testing association close to its full sample magnitude, at 0.564. The vaccination association increases to 0.304, while the stringency association becomes more negative, at −0.203. The stringency relationship is therefore more sensitive in magnitude to monthly coverage and the accompanying change in country composition. Because the restricted samples differ from the full sample, these comparisons cannot isolate the effect of measurement completeness alone. Across the three analytical samples, testing retains the largest absolute rank association with reported incidence, vaccination remains positively associated, and stringency remains negatively associated. The evidence resolves the research questions at the level of pooled observational relationships. It does not identify intervention effects, establish temporal ordering or demonstrate that the associations hold within individual countries after accounting for shared calendar conditions.

6. Discussion:

Higher reported incidence can reflect both greater infection burden and greater capacity to detect it. Table 6 shows that testing intensity has the strongest observed association with recorded cases, while vaccination coverage is positively associated and policy stringency negatively associated. These patterns distinguish diagnostic activity, population coverage and formal restrictions as different dimensions of public health response. Equation 1 aligns their measurement over jointly observed days, making the comparisons temporally consistent within each retained country month. It does not separate transmission from ascertainment. Lipsitch, Swerdlow, and Finelli, 2020, emphasise that understanding epidemic burden requires surveillance capable of identifying infections beyond routinely detected cases. The contribution here is a comparative assessment showing that these response indicators have different relationships with the same surveillance outcome. Reported incidence therefore requires interpretation alongside the processes through which cases become observable. The testing relationship remains positive when the sample is restricted to records classified as tests performed and when monthly coverage requirements increase, as Table 7 demonstrates. This persistence indicates that the observed pattern is not confined to the broadest mixture of testing definitions or the least complete eligible months. However, testing generated by an expanding outbreak and infections discovered through expanded surveillance remain observationally inseparable. No measured mediator distinguishes these explanations. Greenland et al., 2016, caution that interval estimates depend on analytical assumptions and cannot establish an explanation merely because they exclude a null value. The findings consequently strengthen the case for treating diagnostic activity as part of the observation process, while providing no basis for interpreting testing as the dominant causal determinant of transmission.

The positive vaccination association in Table 6 illustrates why aggregate coverage comparisons cannot substitute for estimates of individual protection. Jara et al., 2021, estimated protective vaccination effects through a national cohort in Chile, whereas the present calculation compares country months without individual vaccination histories or adjustment for epidemic timing. The apparent divergence concerns the quantity being estimated rather than direct evidence against protection. Feikin et al., 2022, further show that vaccine effectiveness against infection changes with time since vaccination, which initial protocol coverage alone does not capture. The persistence of the positive association in Table 7 therefore identifies a boundary of coverage based international comparisons: a population can record increasing vaccination coverage alongside substantial reported incidence. This analysis sharpens the distinction between administrative uptake, effective protection and observed epidemic burden without attributing the aggregate relationship to an unmeasured biological or behavioural mechanism. The stringency findings reveal a different interpretive constraint. Its negative association becomes larger in magnitude when monthly coverage is more complete, but the corresponding sample also changes, as Table 7 shows. The evidence therefore supports consistency in direction while leaving the source of the magnitude change unresolved. Soltesz et al., 2020, demonstrate how conclusions about intervention effectiveness can depend on model structure and attribution assumptions. Here, a composite index records formal restrictions without measuring adherence, enforcement or actual contact reduction. Table 5 also documents substantial differences in the distributions of responses and incidence. Together, these results show why comparable index scales do not guarantee comparable institutional implementation. International policy benchmarking should distinguish recorded restrictions from their implementation rather than interpret a modest pooled correlation as a transferable estimate of policy effectiveness.

The international relevance of these findings lies in improving comparison across reporting systems rather than ranking national intervention success. Tables 6 and 7 retain the same directional pattern across the examined samples, but neither comparison identifies differences between advanced and emerging economies or isolates changes within individual countries. Equation 1 provides consistent observation support, while unequal country coverage and differing testing definitions still constrain generalisation. Decision makers should therefore report case trends alongside testing definitions, vaccination coverage and the number of observed days, and assess whether conclusions change under comparable reporting conditions. The theoretical

implication is that public health response and surveillance outcomes must be analysed jointly because institutional action can affect both infection processes and their measurement. The resulting contribution is a transparent account of stable pooled associations and their interpretive boundaries, rather than an unsupported claim that correlation magnitudes reveal causal pathway shares or intervention priorities.

7. Conclusion and Implications:

International comparisons can mistake greater visibility of disease for poorer public health performance. This analysis reveals distinct observational relationships between institutional responses and recorded epidemic burden, with their directions persisting under restrictions addressing reporting definitions and monthly completeness. The contribution lies in examining these relationships on consistent observation days while demonstrating why their magnitudes cannot rank intervention effectiveness. The findings sharpen the theoretical distinction between processes that influence infection and processes that make infections observable. They also show that administrative measures of preventive effort cannot substitute for direct evidence of protection or implementation. For decision makers, the implication is to interpret national performance alongside diagnostic access, reporting conventions and observation coverage before allocating resources or judging institutional effectiveness. Public health agencies should publish comparable measurement definitions, disclose incomplete reporting periods and examine whether apparent differences persist under consistent coverage requirements. Operational dashboards should distinguish recorded burden from the conditions supporting its detection, enabling more defensible comparisons across health systems. These practices can reduce the risk of penalising transparent reporting or overlooking poorly observed populations. The broader contribution is an evidence framework for interpreting international surveillance comparisons, supporting accountable decisions while preserving the distinction between stable associations and demonstrated causal effects.

The unbalanced panel provides opportunities to examine how reporting continuity influences international comparisons. Future analyses should separate variation within countries from differences between countries and account for shared epidemic shocks. Linking surveillance records to diagnostic eligibility, reporting delays, vaccine histories and implementation measures would permit direct assessment of explanations that aggregate indicators cannot distinguish. Alternative completeness thresholds and samples with stable geographical composition could assess selection sensitivity. Designs exploiting defensible policy variation could evaluate intervention effects, while independent infection estimates could clarify ascertainment differences. These extensions would test transferability across health systems and establish which observed relationships reflect measurement, institutional response or transmission.

References:

1. Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317-372. https://doi.org/10.3368/jhr.50.2.317
2. Dagan, N., Barda, N., Kepten, E., Miron, O., Perchik, S., Katz, M. A., Hernán, M. A., Lipsitch, M., Reis, B., &Balicer, R. D. (2021). BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. New England Journal of Medicine, 384(15), 1412-1423. https://doi.org/10.1056/NEJMoa2101765
3. de Winter, J. C. F., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. Psychological Methods, 21(3), 273-290. https://doi.org/10.1037/met0000079
4. Feikin, D. R., Higdon, M. M., Abu-Raddad, L. J., Andrews, N., Araos, R., Goldberg, Y., Groome, M. J., Huppert, A., O'Brien, K. L., Smith, P. G., Wilder-Smith, A., Zeger, S., Deloria Knoll, M., & Patel, M. K. (2022). Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: Results of a systematic review and meta-regression. The Lancet, 399(10328), 924-944. https://doi.org/10.1016/S0140-6736%2822%2900152-0
5. Funk, S., Salathé, M., & Jansen, V. A. A. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: A review. Journal of the Royal Society Interface, 7(50), 1247-1256. https://doi.org/10.1098/rsif.2010.0142
6. Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. European Journal of Epidemiology, 31(4), 337-350. https://doi.org/10.1007/s10654-016-0149-3
7. Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Tatlow, H. (2021). A global panel database of pandemic policies: Oxford COVID-19 Government Response Tracker. Nature Human Behaviour, 5(4), 529-538. https://doi.org/10.1038/s41562-021-01079-8
8. Hasell, J., Mathieu, E., Beltekian, D., Macdonald, B., Giattino, C., Ortiz-Ospina, E., Roser, M., & Ritchie, H. (2020). A cross-country database of COVID-19 testing. Scientific Data, 7(1), Article 345. https://doi.org/10.1038/s41597-020-00688-8
9. Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., Pinior, B., Thurner, S., & Klimek, P. (2020). Ranking the effectiveness of worldwide COVID-19 government interventions. Nature Human Behaviour, 4(12), 1303-1312. https://doi.org/10.1038/s41562-020-01009-0
10. Hughes, R. A., Heron, J., Sterne, J. A. C., & Tilling, K. (2019). Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. International Journal of Epidemiology, 48(4), 1294-1304. https://doi.org/10.1093/ije/dyz032
11. Jara, A., Undurraga, E. A., González, C., Paredes, F., Fontecilla, T., Jara, G., Pizarro, A., Acevedo, J., Leo, K., Leon, F., Sans, C., Leighton, P., Suárez, P., García-Escorza, H., & Araos, R. (2021). Effectiveness of an inactivated SARS-CoV-2 vaccine in Chile. New England Journal of Medicine, 385(10), 875-884. https://doi.org/10.1056/NEJMoa2107715
12. Larremore, D. B., Wilder, B., Lester, E., Shehata, S., Burke, J. M., Hay, J. A., Tambe, M., Mina, M. J., & Parker, R. (2021). Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. Science Advances, 7(1), Article eabd5393. https://doi.org/10.1126/sciadv.abd5393
13. Lipsitch, M., Swerdlow, D. L., & Finelli, L. (2020). Defining the epidemiology of Covid-19: Studies needed. New England Journal of Medicine, 382(13), 1194-1196. https://doi.org/10.1056/NEJMp2002125

14. Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., Giattino, C., &Rodés-Guirao, L. (2021). A global database of COVID-19 vaccinations. Nature Human Behaviour, 5(7), 947-953. https://doi.org/10.1038/s41562-021-01122-8
15. Our World in Data. (2022, June 23). Ending our COVID-19 testing data updates. https://ourworldindata.org/covid-testing-data-archived
16. Our World in Data. (2024). COVID-19 dataset [Archived data set and documentation]. GitHub. https://github.com/owid/covid-19-data/tree/f0d53320b19703cf565940d8269181a8423d7032/public/data
17. Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. Advances in Methods and Practices in Psychological Science, 1(1), 27-42. https://doi.org/10.1177/2515245917745629
18. Soltesz, K., Gustafsson, F., Timpka, T., Jaldén, J., Jidling, C., Heimerson, A., Schön, T. B., Spreco, A., Ekberg, J., Dahlström, Ö., Bagge Carlson, F., Jöud, A., & Bernhardsson, B. (2020). The effect of interventions on COVID-19. Nature, 588(7839), E26-E28. https://doi.org/10.1038/s41586-020-3025-y