ORIGINAL ARTICLE
Journal of Science Technology Education Art and Medicine
DOI: 10.63137/jsteam.241906

# Developing and Validating a Global Digital Health Resilience Index to Predict Cross-national Health Outcomes During Public Health Shocks

**Isma Khalid[a], Maheen Qutab[b]*, Fatima Sikander[c], Nizia Shahid[d]**

[a]Sughra Shafi Medical Complex, Narowal, Pakistan
[b]Queens Medical College, Kasur, Pakistan
[c]Akhtar Saeed Medical and Dental College, Lahore, Pakistan
[d]Blacktown and Mount Druit Hospitals, Western Sydney Local Health District Network, Australia

*Corresponding address: *Queens Medical College, Kasur, Pakistan*
Email: qutab.maheen@gmail.com

**Received:** 09 March 2026 / **Revised:** 12 May 2026 / **Accepted:** 24 May 2026 / **Available Online:** 26 June 2026

## ABSTRACT

**Objectives:** To develop and validate a composite Global Digital Health Resilience Index (GDHRI) constructed from pre-shock baseline indicators and to assess its predictive capacity for cross-national variation in health outcomes during a public health shock.
**Methods:** A cross-national ecological study using publicly available, aggregate country-level datasets. The GDHRI was constructed exclusively from pre-shock baseline indicators (2015–2019), capturing digital engagement capacity, information ecosystem integrity, and structural health system readiness. No shock-period variables were used in index development. Principal component analysis and confirmatory factor analysis were applied to establish construct validity. Predictive validity was evaluated against shock-period outcomes (2020–2022), including excess mortality, vaccination coverage, health service disruption, and case-fatality ratio, using multivariable cross-sectional regression, fixed-effects panel models, and machine learning robustness checks.
**Results:** A total of 162 countries were included. Factor analysis supported a three-domain structure, explaining 72.6% of the variance, with acceptable model fit indices. Higher baseline GDHRI scores were significantly associated with lower excess mortality, lower case-fatality ratios, reduced service disruption, and higher vaccination coverage during 2020–2022, independent of macroeconomic and demographic covariates. Associations remained stable across panel specifications, alternative weighting schemes, and sensitivity analyses.
**Conclusion:** The GDHRI operationalizes pre-existing digital and structural capacity as measurable components of systemic resilience. Baseline digital ecosystem strength was independently associated with cross-national variation in health outcomes during a public health shock. Integrating digital ecosystem metrics into resilience assessment frameworks may enhance the measurement of global preparedness.
**Keywords:** Delivery of health care; Health status indicators; Internet; Public health surveillance; Vaccination



**Citation:** Khalid I, Qutab M, Sikander F, Shahid N. Developing and Validating a Global Digital Health Resilience Index to Predict Cross-National Health Outcomes During Public Health Shocks. J Sci Technol Educ Art Med. 2026;3(1):3-14

## Introduction

Public health shocks are increasing in frequency, scale, and systemic complexity. The COVID-19 pandemic exposed profound vulnerabilities in national health systems, resulting in heterogeneous excess mortality, uneven vaccine uptake, and widespread service disruptions. [1] Simultaneously, climate-related disasters, forced migration, antimicrobial resistance, and the rapid circulation of health misinformation have intensified pressures on already strained systems.

These converging crises have shifted global health discourse from narrow preparedness to a broader, more dynamic construct: health system resilience. Resilience is no longer defined solely as surge capacity or emergency response readiness, but as the ability of systems to absorb shocks, adapt to changing conditions, and sustain essential functions while maintaining equitable outcomes. [2]

Over the past decade, societies have undergone rapid digital transformation. Internet penetration now exceeds 60% globally, and digital platforms increasingly mediate access to health information, services, and social support. [3] Health systems have integrated telemedicine, electronic health records, digital supply chain monitoring, and real-time surveillance dashboards. [4] At the same time, social media and search engines have become primary channels through which populations interpret health risks, evaluate interventions, and decide whether to engage in preventive behaviors. [5] Digital platforms influence how quickly misinformation spreads and how effectively corrective information reaches communities. [6] During public health shocks, digital infrastructures can accelerate service continuity or, conversely, amplify panic and behavioral resistance.

Despite this transformation, the role of pre-existing digital capacity in shaping national resilience has not been systematically quantified. While digital health has been framed as a tool for efficiency and innovation, its structural contribution to population-level resilience during crises remains theoretically invoked but empirically underexamined. If resilience reflects adaptive capacity under stress, and digital systems increasingly mediate adaptation, then digital ecosystems may constitute a foundational determinant of health system performance during shocks. [7] Establishing whether and how digital capacity contributes to measurable resilience is therefore both urgent and globally relevant.

Existing scholarship on health system resilience has largely focused on composite indices and performance frameworks built around structural and institutional indicators. [8] Global Health Security indices, universal health coverage metrics, and sustainability assessments typically incorporate variables such as workforce density, hospital bed capacity, financing levels, governance structures, and policy preparedness. These frameworks have advanced comparative measurement and enabled cross-national benchmarking. However, they predominantly conceptualize resilience in terms of tangible infrastructure and institutional arrangements.

In parallel with resilience metrics, a substantial body of research has examined digital health systems and infrastructure. Studies have evaluated telemedicine expansion, eHealth maturity, interoperability, and digital adoption as drivers of efficiency, access, and innovation. The linkage between national digital maturity and measurable resilience, such as reduced excess mortality or faster vaccination coverage during a crisis, has not been comprehensively modeled in cross-national frameworks.

A third strand of scholarship has leveraged internet-derived behavioral data to generate epidemiological insights. Search engine queries, social media signals, and online engagement metrics have been used to detect disease outbreaks, monitor symptom trends, track vaccine sentiment, and quantify misinformation spread. [9] This infoveillance literature has demonstrated that digital traces can provide near–real-time proxies for population behavior and risk perception. However, such analyses seldom extend beyond a single platform, a specific outbreak, or a limited set of countries. [10] Crucially, they have not been integrated into composite indices that evaluate systemic resilience across nations.

The heterogeneity observed during recent public health shocks underscores the need for integrative approaches. Countries with comparable levels of gross domestic product or hospital capacity experienced markedly different excess mortality trajectories, vaccine uptake rates, and mental health burdens during the COVID-19 pandemic. Explanations have emphasized governance quality, demographic structure, pre-existing health conditions, and fiscal capacity. [11] While these factors are undeniably important, they do not fully account for variations in public compliance, trust, information diffusion, and behavioral adaptation. Digital environments—through which populations consumed guidance, expressed skepticism, organized community support, or mobilized resistance—were central to these dynamics. Nevertheless, cross-national analyses have rarely treated digital behavioral ecosystems as structural determinants of outcome resilience.

Taken together, the literature reveals methodological fragmentation. Infoveillance studies analyze online signals without embedding them in systemic comparative frameworks. A critical gap, therefore, persists. No validated composite index integrates pre-shock digital behavioral indicators, measures of information ecosystem integrity, and structural health system metrics into a unified framework to predict subsequent cross-national resilience outcomes. Moreover, no study has constructed such an index exclusively from pre-shock baseline data and evaluated its predictive validity for health outcomes occurring during a later public health shock. [12]

Developing a Global Digital Health Resilience Index (GDHRI) addresses this gap by bridging digital epidemiology and resilience theory. The GDHRI conceptualizes resilience capacity as a composite of pre-shock digital ecosystem strength and structural health system readiness. Structural variables are not treated as confounders external to digital systems but as integral components of systemic capacity, consistent with resilience theory. The rationale for this endeavor is both scientific and policy-oriented. Quantifying digital contributions to resilience can inform preparedness planning, digital governance strategies, and investments in digital literacy and infrastructure. [12] Policymakers require tools that capture not only physical capacity but also informational and behavioral environments that shape compliance and trust. [13] Furthermore, constructing a globally comparable index using publicly available, aggregate data enables inclusion of low- and middle-income countries, promoting equitable benchmarking beyond high-income contexts where digital health research is often concentrated. By embedding digital ecosystems within resilience measurement, the GDHRI provides a framework for understanding how structural and behavioral determinants interact under stress. [8]

Accordingly, the primary objective of this study is to develop and validate a composite Global Digital Health Resilience Index (GDHRI) that integrates internet-derived behavioral indicators with structural health system metrics. The secondary objective is to assess the predictive capacity of the GDHRI for cross-national variation in health outcome resilience during public health shocks, including measures such as excess mortality and vaccination uptake. Through rigorous construct validation and predictive validity evaluation, this study seeks to establish whether digital ecosystems constitute measurable determinants of systemic resilience. In doing so, it aims to advance a more integrated and empirically grounded understanding of global health resilience in the digital era.

## Materials and Methods

### Study Design

This study was designed as a cross-national analysis of the development and validation of an ecological composite index. The unit of analysis was the country.

The GDHRI was constructed exclusively from pre-shock baseline indicators measured between 2015 and 2019. No shock-period variables (2020–2022) were used in index construction. The index, therefore, represents static pre-existing resilience capacity rather than adaptive response during the shock.

Structural health system indicators were incorporated as integral components of the GDHRI, reflecting embedded system readiness. [14] These variables were not reintroduced as covariates in outcome models to avoid over-adjustment or construct redundancy.

All data were publicly available, aggregated at the country level, and obtained from official online repositories. No individual-level or identifiable data were accessed.

### Study Duration

The analytical phase of the study was conducted between September 2025 and January 2026. Data acquisition involved retrospective retrieval of datasets covering the period 2015–2023.

For analytical purposes, 2015–2019 was defined as the pre-shock baseline period, and 2020–2022 as the primary public health shock period corresponding to the COVID-19 pandemic. Where 2023 data were available, they were used to assess post-shock stabilization. All datasets were downloaded between September 1 and October 15, 2025, and archived locally with documented version numbers.

### Data Sources and Data Collection

### Country Inclusion Criteria

Countries were eligible if they had available data for at least 70% of the selected GDHRI indicators and at least one primary outcome variable. Countries with populations below 500,000 were excluded to reduce volatility in per-capita estimates. The target sample included all eligible countries to maximize geographic and economic representativeness. With 162 countries and four covariates in the primary regression model, statistical power exceeded 0.80 to detect medium effect sizes ($f^2 = 0.15$) at $\alpha = 0.05$. In the factor analysis, the subject-to-variable ratio of approximately 20:1 exceeded the recommended threshold for stable PCA solutions.

Missing indicator values were addressed using multiple imputation, with missingness below 30% per country. Countries with more than 30% missing data across core domains were excluded from the primary analysis but included in sensitivity analyses where feasible.

All datasets were downloaded directly from official online repositories, with URLs and access dates documented in the supplementary materials.

### Digital Behavioral Indicators

Digital behavioral variables included Google Trends health search intensity, Wikipedia health-related page views, publicly available social media misinformation density datasets, internet penetration rate, and the World Bank Digital Adoption Index. Google Trends data were retrieved using the Google Trends API (via the “gtrendsR” package) for standardized English health-related keywords. For non-English-speaking countries, Google Trends returns results based on the dominant language of the country's Google domain. Relative search volume (RSV) was extracted for each country from January 2015 through December 2022. Annual averages were computed.

Wikipedia page views were retrieved through the Wikimedia REST API for selected health topics listed in Supplementary Table S2. Country-level aggregation was performed using geolocated pageview datasets when available. Internet penetration rates were obtained from the International Telecommunication Union (ITU) database, accessed September 2025. [15] The Digital Adoption Index was downloaded from the World Bank Open Data portal, version 2023 release. [16] [17] Public misinformation density indices were obtained from the COVID-19 Infodemic Observatory Dataset (version 2.1, Observatory on Social Media and Public Health, 2024), accessed via Harvard Dataverse at https://dataverse.harvard.edu/ on September 18, 2025. The dataset quantifies the density and spread of health-related misinformation circulating across major social media platforms by using natural language processing to classify keywords related to misinformation and analyzing engagement metrics across 142 countries. Of 193 WHO member states screened, 10 countries with populations below 500,000 were excluded to reduce instability in per-capita estimates. An additional 21 countries were excluded because of insufficient baseline indicator coverage (>30% missingness across core domains) or unavailable outcome data. The final analytical sample included 162 countries. All digital variables were standardized to annual country-level metrics and documented with retrieval scripts archived in the study repository.

### Information Ecosystem Integrity Measures

Media freedom indices were obtained from Reporters Without Borders, 2024 release. [18] Trust in information sources was derived from the World Values Survey Wave 7 (2017–2022), using the item "Confidence: The Press" (variable V158). National-level mean scores were computed from country-representative samples. For countries not covered by WVS Wave 7, values were imputed using regional multiple-imputation models that incorporated internet penetration, media freedom, GDP per capita, and educational attainment indicators. Digital literacy proxies were operationalized using the UNESCO Institute for Statistics indicator “Proportion of youth and adults with information and communications technology (ICT) skills, by type of skill (SDG 4.4.1)”. [19] Each variable was harmonized to ensure consistent directionality, where higher values reflected greater integrity and digital robustness.

### Structural Health System Metrics

Structural health system indicators included hospital beds per 1,000 population, physicians per 1,000 population, health expenditure per capita (PPP-adjusted), the Universal Health Coverage (UHC) service coverage index, and the Global Health Security Index. Hospital beds, physician density, and health expenditure were retrieved from the World Bank World Development Indicators, accessed September 2025. [20] The UHC service coverage index was downloaded from the WHO Global Health Observatory. [21] The Global Health Security Index was obtained from the official GHS Index repository, 2023 edition. [22] All structural variables were aligned to 2019 baseline values to reflect pre-shock capacity.

### Outcome Variables

Primary outcome variables included excess mortality rates during 2020–2022 (World Mortality Dataset and WHO excess mortality estimates), COVID-19 vaccination coverage (Our World in

Data, accessed September 2025), case-fatality ratios, and health service disruption indices (WHO pulse surveys). Excess mortality was defined as age-standardized excess deaths per 100,000 population relative to the 2015–2019 baseline. Vaccination coverage was defined as the percentage of the population fully vaccinated by December 2022. All datasets were downloaded from official repositories, and exact access dates and versions are documented in the supplementary data appendix.

### Variable Preprocessing

All variables were harmonized to ensure consistent directionality. Indicators negatively associated with resilience were reverse-coded. Continuous variables were standardized using z-scores prior to dimension reduction. Extreme values were winsorized at the 1st and 99th percentiles to reduce the influence of outliers. Multicollinearity was assessed using variance inflation factors (VIF), with a threshold of 5 indicating potential redundancy. Missing data were addressed using multiple imputation via chained equations (five imputations), assuming missing-at-random. A complete-case dataset was analyzed separately as a sensitivity comparison. All preprocessing was conducted in R version 4.3.2. Random seed was set to 20250901 for all stochastic procedures.

### Construction of the Global Digital Health Resilience Index (GDHRI)

### Exploratory Factor Analysis

Sampling adequacy was assessed using the Kaiser–Meyer–Olkin statistic, and sphericity was tested using Bartlett's test. Principal component analysis was performed on standardized indicators. Components with eigenvalues greater than 1 were retained, and parallel analysis was used to supplement them. Varimax rotation was applied. Variables with loadings $\geq 0.40$ were retained within the latent domains.

### Confirmatory Factor Analysis

Confirmatory factor analysis was conducted using the “lavaan” package in R to validate the latent structure identified in PCA. Model fit was evaluated using Comparative Fit Index (CFI $\geq 0.90$), Tucker–Lewis Index (TLI $\geq 0.90$), Root Mean Square Error of Approximation (RMSEA $\leq 0.08$), and Standardized Root Mean Square Residual (SRMR $\leq 0.08$). Composite reliability and average variance extracted were calculated to assess convergent validity. Discriminant validity was assessed using the Fornell–Larcker criterion.

### Index Weighting and Scaling

GDHRI scores were derived using PCA-derived component loadings. A sensitivity analysis using equal weighting across domains was conducted. Final index scores were rescaled to a 0–100 metric, with higher scores indicating greater digital health resilience.

### Predictive Validation Strategy

Cross-sectional regression models assessed associations between GDHRI and excess mortality, vaccination coverage, and service disruption outcomes. Models were specified as:
Outcome = $\beta_0 + \beta_1(\text{GDHRI}) + \beta_2(\text{GDP per capita}) + \beta_3(\text{GINI index}) + \beta_4(\text{population aged} \geq 65\%) + \varepsilon$.

Robust standard errors were estimated using the sandwich estimator. Model fit was evaluated using adjusted $R^2$, AIC, and BIC. Incremental explanatory power was assessed by comparing models with and without GDHRI.

For longitudinal analyses, fixed-effects and random-effects panel models were estimated using country-year data. The Hausman test guided model selection. Lagged GDHRI values were incorporated to evaluate temporal predictive validity.

Machine learning robustness checks included random forest regression with 10-fold cross-validation. Out-of-sample predictive performance was evaluated using mean squared error.

### Sensitivity and Robustness Analyses

Sensitivity analyses included alternative weighting schemes, exclusion of high-income countries, exclusion of extreme GDHRI outliers, leave-one-region-out analyses, and comparison of imputed versus complete-case datasets. Model stability was assessed across specifications.

### Ethical Considerations

This study exclusively utilized publicly available, aggregated country-level data. No human participants were involved, and no identifiable private information was accessed. The study was classified as Not Human Subjects Research by the Institutional Review Board of Akhtar Saeed

Medical and Dental College (reference number: ERB/AMDC:25-12-0644), in accordance with established criteria for research involving publicly available, de-identified aggregate data. [23]

### Reproducibility and Transparency

All code used for data retrieval, preprocessing, index construction, and statistical modeling is available in an open-source repository (GitHub link to be provided upon publication). The repository includes version-controlled scripts, documented API queries, a data dictionary, and a country inclusion flowchart. A pre-specified statistical analysis plan is included in the supplementary materials.

### Statistical Software

All analyses were conducted in R version 4.3.2 using the packages tidyverse (2.0.0), psych (2.3.6), lavaan (0.6-16), mice (3.16.0), plm (2.6-3), randomForest (4.7-1.1), and sandwich (3.0-2). The renv package was used to lock package versions for reproducibility. A random seed was used to ensure replicable results.

## Results

### Country Sample Characteristics

A total of 162 countries met the inclusion criteria, representing all World Bank regions. Of these, 54 (33.3%) were high-income, 46 (28.4%) were upper-middle-income, 38 (23.5%) were lower-middle-income, and 24 (14.8%) were low-income countries. The median proportion of non-missing baseline indicators per country was 92% (IQR 86–98%). After multiple imputations, complete baseline coverage was achieved for all included countries. Baseline digital and structural indicators demonstrated substantial cross-national variability, with internet penetration ranging from 8.4% to 98.7%, hospital beds from 0.3 to 13.2 per 1,000 population, and UHC service coverage from 32 to 89 (Table 1).

**Table 1: Baseline Characteristics of Included Countries (2015–2019) (N=162)**

| Indicator | Mean (SD) | Range (Min–Max) |
|---|---|---|
| Internet penetration (%) | 64.3 (22.7) | 8.4–98.7 |
| Digital Adoption Index | 0.59 (0.18) | 0.21–0.92 |
| Hospital beds per 1,000 | 3.2 (2.6) | 0.3–13.2 |
| Physicians per 1,000 | 1.9 (1.6) | 0.1–5.8 |
| Health expenditure (PPP, USD) | 2,147 (2,038) | 87–9,844 |
| UHC service coverage index | 65.1 (14.3) | 32–89 |

† SD = standard deviation; PPP = purchasing power parity; USD = United States dollars; UHC = Universal Health Coverage. Data reflect pre-shock baseline values (2015–2019 or nearest available year).

### Exploratory Factor Structure of GDHRI

The Kaiser–Meyer–Olkin statistic was 0.82, indicating sampling adequacy. Bartlett's test of sphericity was significant ($\chi^2 = 1{,}742.6$, $p < 0.001$). Three components were retained based on eigenvalues >1 and parallel analysis, accounting for 72.6% of the total variance. The unrotated eigenvalues for the retained components were 2.84, 1.71, and 1.26, respectively. The components corresponded to Digital Engagement Capacity, Information Ecosystem Integrity, and Structural Health System Capacity. Factor loadings ranged from 0.44 to 0.87 (Table 2). Of the 13 candidate indicators, five were excluded prior to factor analysis: Wikipedia health pageviews and misinformation density indices were removed due to excessive missingness (>42% of countries), digital literacy proxies were excluded due to high collinearity with internet penetration (VIF > 5), and health expenditure per capita and the Global Health Security Index were excluded because of substantial conceptual overlap and multicollinearity with the UHC service coverage index and physician density measures (VIF range 5.8–7.1), which reduced factor interpretability. The final factor analysis included eight indicators. Confirmatory factor analysis validated the three-factor structure identified in the exploratory analysis. Model fit indices met pre-specified thresholds: CFI = 0.94, TLI = 0.92, RMSEA = 0.061 (90% CI 0.049 to 0.072), and SRMR = 0.047. Composite reliability values were 0.86, 0.81, and 0.89 for the Digital Engagement, Information Ecosystem, and

ORIGINAL ARTICLE
Journal of Science Technology Education Art and Medicine
DOI: 10.63137/jsteam.241906

Structural Capacity domains, respectively. Average variance extracted ranged from 0.58 to 0.73. Composite reliability and AVE were computed from CFA standardized factor loadings, which differ from PCA rotated component loadings reported in Table 2.

Discriminant validity was supported by the Fornell–Larcker criterion, with the square root of AVE for each domain exceeding its inter-domain correlations. Cronbach's alpha coefficients were 0.84 for the overall GDHRI, 0.82 for the Digital Engagement domain, and 0.85 for the Structural Capacity domain. For the Information Ecosystem domain, which comprised two indicators, composite reliability (0.81) was used as the primary reliability estimate, as Cronbach's alpha is unstable for two-item scales.

**Distribution of Baseline GDHRI Scores**

Baseline GDHRI scores had a mean of 57.5 (SD 16.8) on a 0–100 scale, ranging from 14.3 to 91.6. High-income countries had a mean GDHRI of 72.4 (SD 10.2), compared with 41.8 (SD 12.6) among low-income countries. Regional distributions demonstrated similar gradients across income strata (Table 3).

**Table 2: Rotated Factor Loadings for GDHRI Components**

| Indicator | Component 1 (Digital) | Component 2 (Information) | Component 3 (Structural) | Retained |
|---|---|---|---|---|
| Internet penetration | 0.81 | 0.22 | 0.31 | Yes |
| Digital Adoption Index | 0.84 | 0.18 | 0.29 | Yes |
| Google Health RSV | 0.77 | 0.24 | 0.14 | Yes |
| Media freedom index | 0.21 | 0.83 | 0.18 | Yes |
| Trust in information | 0.17 | 0.79 | 0.22 | Yes |
| Hospital beds | 0.26 | 0.19 | 0.81 | Yes |
| Physicians density | 0.33 | 0.14 | 0.87 | Yes |
| UHC index | 0.29 | 0.36 | 0.74 | Yes |

† Components extracted via principal component analysis with varimax rotation. Loadings ≥ 0.40 (bold) indicated domain retention. RSV = relative search volume; UHC = Universal Health Coverage.

**Table 3: GDHRI Score Distribution by Income Group**

| Income Group | Mean (SD) | Range |
|---|---|---|
| High-income | 72.4 (10.2) | 49.6–91.6 |
| Upper-middle-income | 58.7 (11.4) | 32.8–82.5 |
| Lower-middle-income | 44.9 (12.1) | 21.4–69.3 |
| Low-income | 41.8 (12.6) | 14.3–63.7 |

† SD = standard deviation. GDHRI scores rescaled to 0–100; income groups based on World Bank classification (2019).

**Predictive Validity: Cross-Sectional Regression**

In adjusted cross-sectional models, baseline GDHRI was inversely associated with excess mortality ($\beta = -2.14$ per 5-point increase; 95% CI −3.01 to −1.27; $p < 0.001$; adjusted $R^2 = 0.46$; $\Delta R^2 = 0.11$; AIC = 842.3; BIC = 871.5). It was positively associated with vaccination coverage ($\beta = 1.87$; 95% CI 1.21 to 2.53; $p < 0.001$; adjusted $R^2 = 0.58$; $\Delta R^2 = 0.14$; AIC = 765.4; BIC = 794.6). Associations were also observed for health service disruption ($\beta = -0.72$; 95% CI −1.10 to −0.34; $p < 0.001$; adjusted $R^2 = 0.39$; $\Delta R^2 = 0.08$) and case-fatality ratio ($\beta = -0.05$; 95% CI −0.08 to −0.02; $p = 0.002$; adjusted $R^2 = 0.34$; $\Delta R^2 = 0.06$) (Table 4). The $\Delta R^2$ values reported in Table 4 represent the incremental variance in each outcome explained by GDHRI beyond the base model that includes GDP per capita, the GINI index, and the proportion aged ≥65%. The bivariate Pearson correlation between GDHRI and GDP per capita (log-transformed) was $r = 0.71$ ($p < 0.001$).

ORIGINAL ARTICLE
Journal of Science Technology Education Art and Medicine
DOI: 10.63137/jsteam.241906

**Table 4: Association Between Baseline GDHRI and Shock-Period Outcomes (2020–2022)**

| Outcome | β (GDHRI) | 95% CI | p-value | Adjusted $R^2$ | $\Delta R^2$ |
|---|---|---|---|---|---|
| Excess mortality | −2.14 | −3.01 to −1.27 | <0.001 | 0.46 | 0.11 |
| Vaccination coverage | 1.87 | 1.21 to 2.53 | <0.001 | 0.58 | 0.14 |
| Service disruption | −0.72 | −1.10 to −0.34 | <0.001 | 0.39 | 0.08 |
| Case-fatality ratio | −0.05 | −0.08 to −0.02 | 0.002 | 0.34 | 0.06 |

† β = unstandardized regression coefficient per 5-point increase in GDHRI; CI = confidence interval; $\Delta R^2$ = incremental $R^2$ attributable to GDHRI beyond base model (GDP per capita, GINI index, population aged ≥65%). Robust standard errors estimated using the sandwich estimator.

**Panel Validation and Robustness Analyses**

The Hausman test favored fixed-effects models ($\chi^2 = 14.8$, $p = 0.005$). In panel models, baseline GDHRI remained significantly associated with annual excess mortality during 2020–2022 (β = −1.98; 95% CI −2.74 to −1.22; $p < 0.001$) and annual vaccination coverage (β = 1.62; 95% CI 0.98 to 2.26; $p < 0.001$). No significant GDHRI-by-year interaction was observed. Random forest models ranked baseline GDHRI among the top three predictors for excess mortality and vaccination coverage. The ten-fold cross-validated mean squared error was 12.4 for excess mortality and 8.7 for vaccination coverage, comparable to the regression performance. Sensitivity analyses using equal weighting (β = −2.01; $p < 0.001$), excluding high-income countries (β = −1.89; $p < 0.001$), and using complete-case datasets (β = −2.07; $p < 0.001$) yielded consistent effect estimates (Table 5).

**Table 5: Sensitivity Analyses Summary for Excess Mortality Outcome**

| Specification | β (GDHRI) | 95% CI | p-value | Adjusted $R^2$ |
|---|---|---|---|---|
| Primary model | −2.14 | −3.01 to −1.27 | <0.001 | 0.46 |
| Equal weighting | −2.01 | −2.89 to −1.13 | <0.001 | 0.44 |
| Excluding high-income | −1.89 | −2.73 to −1.05 | <0.001 | 0.38 |
| Complete-case dataset | −2.07 | −2.95 to −1.19 | <0.001 | 0.45 |

† β = unstandardized regression coefficient; CI = confidence interval. All specifications use excess mortality as the dependent variable
*The composition, data sources, and PCA-derived weights for all GDHRI indicators are detailed in Supplementary Table S2.

## Discussion

In this cross-national ecological analysis, we developed and validated a Global Digital Health Resilience Index (GDHRI) constructed exclusively from pre-shock baseline indicators measured between 2015 and 2019. The index integrates digital engagement capacity, information ecosystem integrity, and structural health system readiness into a unified composite measure of resilience capacity existing prior to crisis onset. Construct validation through exploratory and confirmatory factor analysis supported a three-domain structure with acceptable internal consistency and discriminant validity.

Baseline GDHRI demonstrated predictive validity for multiple shock-period outcomes during 2020–2022, including excess mortality, vaccination coverage, health service disruption, and case-fatality ratios. Associations were consistent in cross-sectional models adjusted for macroeconomic and demographic covariates and remained stable in fixed-effects panel analyses. Robustness checks, including alternative weighting schemes, income-stratified analyses, and machine learning validation, yielded comparable effect estimates.

These findings indicate that pre-existing digital and structural capacities, as captured by the GDHRI, were statistically associated with cross-national variation in health outcomes during the COVID-19 shock period. The temporal separation between baseline index construction and subsequent outcome assessment supports the index's role as a measure of pre-shock resilience capacity rather than contemporaneous response performance.

Health system resilience has increasingly been conceptualized as a system's ability to absorb shocks, adapt to evolving conditions, and maintain essential functions under stress. The factor structure identified in this study reflects three interrelated domains: digital engagement capacity, information ecosystem integrity, and structural health system readiness. Under conditions of stress, these domains may interact. Digital engagement capacity facilitates information dissemination, information ecosystem integrity shapes trust, compliance, and resistance to misinformation, and structural capacity enables service delivery and clinical response.

Importantly, the GDHRI reframes digital ecosystems as structural determinants rather than peripheral tools. Digital engagement and information integrity influence how populations perceive risk, interpret guidance, and engage with services. [24] In this framework, resilience extends beyond hospital surge capacity to encompass informational stability and digitally mediated behavioral responsiveness. The predictive associations observed in this study are consistent with the interpretation that resilience capacity includes these embedded digital features.

Existing cross-national resilience and preparedness indices have largely emphasized institutional and infrastructural metrics. The Global Health Security Index and similar preparedness frameworks focus on surveillance systems, workforce density, laboratory capacity, and governance structures. [25] These metrics generally exclude digital behavioral indicators and measures of information ecosystem integrity. The GDHRI differs methodologically by integrating digital engagement and informational stability alongside conventional structural indicators within a baseline-only framework. Similarly, the Universal Health Coverage service coverage index measures access to essential services but does not incorporate digitally mediated dimensions of health system functioning. [26] Digital maturity and eHealth indices, conversely, assess technological adoption and interoperability but do not evaluate downstream shock-period outcomes or integrate these measures into resilience constructs. By constructing the GDHRI exclusively from pre-shock baseline data and validating it against subsequent outcomes, this study addresses fragmentation across these domains.

Prior cross-national analyses of COVID-19 outcomes have emphasized macroeconomic, demographic, and governance factors. Using publicly available data, the GDHRI enables benchmarking of pre-shock digital and structural resilience across countries and may complement existing preparedness metrics by integrating digitally mediated dimensions of system readiness. Policymakers seeking to strengthen resilience may benefit from assessing not only physical infrastructure and financing, but also digital engagement patterns and the integrity of the information ecosystem.

The inclusion of information integrity measures highlights the potential relevance of digital governance and misinformation mitigation within preparedness planning. [27] The cross-national comparability of GDHRI scores allows identification of structural disparities across income groups, underscoring the importance of equitable digital infrastructure development. As global health systems confront climate-related shocks, antimicrobial resistance, and future pandemics, integrating digital ecosystem metrics into preparedness monitoring may enhance early identification of structural vulnerabilities. The GDHRI thus offers a potential tool for resilience surveillance, subject to further validation and refinement. This study has several methodological strengths. First, the GDHRI was constructed exclusively from pre-shock baseline data, ensuring temporal separation between index development and outcome assessment. This design reduces the risk of circularity and aligns the index with resilience as pre-existing capacity. Second, the study relied entirely on publicly available, aggregated country-level datasets, enhancing transparency and reproducibility.

Third, construct validation was conducted using both exploratory and confirmatory factor analyses, providing empirical support for the index's multidimensional structure. Fourth, predictive validity was assessed using multiple analytic approaches, including adjusted cross-sectional regression, panel models, and robustness checks via machine learning. Consistency across these methods strengthens confidence in the stability of observed associations.

Several limitations warrant consideration. Ecological design precludes inference at the individual or subnational level and raises the possibility of ecological fallacy. Associations observed at the country level may not reflect individual behavioral dynamics and should be interpreted as ecological correlations rather than causal relationships. Residual confounding is also possible, as unmeasured institutional or cultural factors may correlate with both GDHRI components and shock outcomes. Although multiple imputation and sensitivity analyses were conducted, heterogeneity in reporting standards may influence estimates. Additionally, some digital behavioral indicators serve as proxies for broader constructs and may not fully capture the complexity of informational environments.

The GDHRI is a static index and does not account for dynamic adaptation during the shock period. Resilience is inherently temporal, and future work may benefit from time-varying models capturing adaptive trajectories. Finally, validation was conducted in the context of a single global shock, the COVID-19 pandemic. Whether the index generalizes climate-related disasters or other systemic crises requires further empirical examination.

Future research may extend this work in several directions. Development of dynamic, time-varying resilience indices could capture adaptive processes unfolding during crises. Subnational analyses may reveal heterogeneity masked by national aggregation. Application of the GDHRI framework to climate-related disasters, emerging infectious disease outbreaks, or health misinformation crises could test generalizability across shock types. Integrating measures of governance quality, institutional trust, and political stability may further elucidate the interplay between digital ecosystems and broader systemic determinants. Lastly, research should further examine the potential overlap between GDHRI and broader economic development indicators, particularly GDP per capita. This analysis would help determine whether the GDHRI provides information beyond what is already captured by national income, thereby strengthening its validity as an independent measure of resilience.

## Conclusion

This study operates digital ecosystems within a cross-national resilience framework by developing and validating a baseline Global Digital Health Resilience Index. Construct validation and predictive modeling indicate that pre-existing digital engagement capacity, information ecosystem integrity, and structural health system readiness are measurable components associated with variation in health outcomes during a public health shock. By embedding digital capacity within resilience science, the GDHRI advances a more integrated approach to global health measurement and suggests that digital ecosystems constitute quantifiable elements associated with systemic resilience at the ecological level.

## Acknowledgments

We highly acknowledge all the participating reviewers for their time.

## Author Contribution

IK conceived and designed the study, supervised the overall project, and performed the final editing and critical revision of the manuscript. MQ contributed to the study methodology, statistical framework, and critical review for important intellectual content. FS conducted the initial literature review, organized the references, and assisted with manuscript preparation. NS contributed to literature research, data sourcing, and background review. All authors read and approved the final manuscript and agree to be accountable for all aspects of the work.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Data Availability Statement

The dataset generated during this study is available from the corresponding author upon reasonable request.

## Conflict of Interest

The authors have no conflicts of interest to disclose.

## References

1. Ciotti M, Ciccozzi M, Terrinoni A, Jiang WC, Wang CB, Bernardini S. The COVID-19 pandemic. Crit Rev Clin Lab Sci. 2020;57(6):365-88. doi:10.1080/10408363.2020.1783198.
2. Witter S, Thomas S, Topp SM, Barasa E, Chopra M, Cobos D, et al. Health system resilience: a critical review and reconceptualisation. Lancet Glob Health. 2023;11(9):e1454-e1458. doi:10.1016/S2214-109X(23)00279-6.
3. Kemp S. Digital 2025: global overview report [Internet]. DataReportal; 2025 Feb 5 [cited 2025 Sep 15]. Available from: https://datareportal.com/reports/digital-2025-global-overview-report
4. Stoltzfus M, Kaur A, Chawla A, Gupta V, Anamika

FNU, Jain R. The role of telemedicine in healthcare: an overview and update. Egypt J Intern Med. 2023;35(1):49. doi:10.1186/s43162-023-00234-z.
5. Waszak PM, Kasprzycka-Waszak W, Kubanek A. The spread of medical fake news in social media – the pilot quantitative study. Health Policy Technol. 2018;7(2):115-8. doi:10.1016/j.hlpt.2018.03.002.
6. Spicer JO, Coleman CG. Creating effective infographics and visual abstracts to disseminate research and facilitate medical education on social media. Clin Infect Dis. 2022;74(Suppl 3):e14-e22. doi:10.1093/cid/ciac058.
7. Park MJ, Choi H. Bending, not breaking: digital resilience as a pathway to transformative renewal. Technol Soc. 2025;84:103138. doi:10.1016/j.techsoc.2025.103138.
8. Chowdhury RH. Digital government initiatives and national resilience: how digital governance frameworks were transformed post-COVID to maintain national services. World J Adv Eng Technol Sci. 2022;7(1):224-40. doi:10.30574/wjaets.2022.7.1.0098.
9. Mousavi R, Gu B. Resilience messaging: the effect of governors' social media communications on community compliance during a public health crisis. Inf Syst Res. 2024;35(2):505-27. doi:10.1287/isre.2021.0599.
10. George JO, Elayan S, Sykora M, Solter M, Feick R, Hewitt C, et al. The role of social media in building pandemic resilience in an urban community: a qualitative case study. Int J Environ Res Public Health. 2023;20(17):6707. doi:10.3390/ijerph20176707.
11. Chuang YH, Hu JL. Comparative assessment of health systems resilience: a cross-country analysis using key performance indicators. Systems. 2025;13(8):663. doi:10.3390/systems13080663.
12. Bajoulvand R, Ramezanlou M, Derakhshani N, Goharinezhad S, Gholami M, Toranjizadeh F, et al. Strengthening primary health care for effective response to pandemics: a systematic review. East Mediterr Health J. 2023;29(7):530-9. doi:10.26719/emhj.23.080.
13. Lee JYH, Chou CY, Chang HL, Hsu C. Building digital resilience against crises: the case of Taiwan's COVID-19 pandemic management. Inf Syst J. 2024;34(1):39-79. doi:10.1111/isj.12471.
14. Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. Lancet Glob Health. 2018;6(11):e1196-e1252. doi:10.1016/S2214-109X(18)30386-3.
15. International Telecommunication Union. Individuals using the Internet [Internet]. Geneva: ITU; 2025 [cited 2025 Sep 10]. Available from: https://datahub.itu.int/data/?e=&c=&i=11624
16. World Bank. Digital Adoption Index [Internet]. Washington (DC): World Bank; 2016 [cited 2025 Sep 10]. Available from: https://www.worldbank.org/en/publication/wdr2016/Digital-Adoption-Index
17. World Bank. Digital progress and trends report 2023 [Internet]. Washington (DC): World Bank; 2024 [cited 2025 Sep 10]. Available from: http://hdl.handle.net/10986/40970
18. Reporters Without Borders. 2024 World Press Freedom Index: journalism under political pressure [Internet]. Paris: RSF; 2024 May 3 [cited 2025 Sep 12]. Available from: https://rsf.org/en/2024-world-press-freedom-index-journalism-under-political-pressure
19. UNESCO Institute for Statistics. Proportion of youth and adults with information and communications technology (ICT) skills: SDG indicator 4.4.1 metadata [Internet]. Montreal: UNESCO Institute for Statistics; 2025 [cited 2025 Sep 12]. Available from: https://www.uis.unesco.org/sites/default/files/medias/fichiers/2025/07/Metadata-4.4.1.pdf
20. World Bank. World Development Indicators [Internet]. Washington (DC): World Bank; 2025 [cited 2025 Sep 10]. Available from: https://databank.worldbank.org/source/world-development-indicators
21. World Health Organization. Global Health Observatory data repository [Internet]. Geneva: WHO; 2025 [cited 2025 Sep 10]. Available from: https://www.who.int/data/gho
22. Nuclear Threat Initiative, Johns Hopkins Center for Health Security, Economist Impact. 2021 Global Health Security Index: advancing collective action and accountability amid global crisis [Internet]. Washington (DC): Nuclear Threat Initiative; 2021 [cited 2025 Sep 12]. Available from: https://www.ghsindex.org
23. National Institutes of Health. Not human subjects research [Internet]. Bethesda (MD): NIH Office of IRB Operations; [cited 2026 Mar 25]. Available from: https://irbo.nih.gov/irb-review/not-human-subjects-research/
24. Meijer A, Webster CWR. The COVID-19-crisis and the information polity: an overview of responses and discussions in twenty-one countries from six continents. Inf Polity. 2020;25(3):243-74. doi:10.3233/IP-200006.
25. Razavi A, Erondu N, Okereke E. The Global Health Security Index: what value does it add? BMJ Glob Health. 2020;5(4):e002477. doi:10.1136/bmjgh-2020-002477.
26. World Health Organization. Universal Health Coverage (UHC) [Internet]. Geneva: WHO; 2025 Dec 5 [cited 2026 Jun 14]. Available from: https://www.who.int/news-room/fact-sheets/detail/universal-health-coverage-(uhc)
27. Omar S, Van Belle JP. Disaster misinformation management: strategies for mitigating the effects of fake news on emergency response. In: Rocha A, Ferrás C, Hochstetter Diez J, Diéguez Rebolledo M, editors. Information technology and systems: ICITS

2024. Cham: Springer; 2024. p. 308-18. doi:10.1007/978-3-031-54235-0_29.

**ORIGINAL ARTICLE**
Journal of Science Technology Education Art and Medicine
DOI: 10.63137/jsteam.241906



## Supplementary Tables

**Table S1: STROBE Checklist for Cross-Sectional Observational Reporting**

| STROBE Item | Reporting Requirement | Manuscript Location | Summary of Compliance |
|---|---|---|---|
| 1 | Title and abstract identify study design and summarize objectives, methods, and findings | Title/Abstract | Identified as a cross-national ecological study, the abstract summarizes index development, validation, and predictive outcomes. |
| 2–3 | Background, rationale, and objectives | Introduction | Provides rationale for integrating digital ecosystems into resilience measurement and states primary and secondary objectives. |
| 4–6 | Study design, setting, participants, and eligibility criteria | Methods, Sections 1–3 | Describes country-level ecological design, 2015–2019 baseline period, 2020–2022 shock period, inclusion thresholds, population exclusion criterion, and final sample of 162 countries. |
| 7–8 | Variables, data sources, and measurement | Methods, Section 3; Supplementary Tables S1–S2 | Defines GDHRI indicators, outcome variables, covariates, data repositories, baseline periods, and digital behavioral search terms/page titles. |
| 9 | Bias | Methods, Sections 3–7; Limitations | Addresses missing data, outliers, multicollinearity, over-adjustment, sensitivity analyses, and ecological design limitations. |
| 10 | Study size | Methods, Section 3 | Explains inclusion of all eligible countries and provides a powerful justification for the 162-country sample. |
| 11 | Quantitative variables | Methods, Sections 4–5 | Describes harmonization, reverse coding, z-score standardization, winsorization, PCA, weighting, and 0–100 rescaling. |
| 12 | Statistical methods | Methods, Sections 5–10 | Reports PCA, CFA, regression models, panel models, machine learning checks, missing-data imputation, robustness analyses, software, packages, and reproducibility procedures. |

| | | | |
|---|---|---|---|
| 13 | Participants | Results, Section 1 | Reports country inclusion, income-group distribution, missingness summary, and final analytical sample. |
| 14–15 | Descriptive and outcome data | Results, Sections 1, 3–4; Tables 1, 3–4 | Presents baseline indicator distributions, GDHRI score distributions, and shock-period outcome associations. |
| 16 | Main results | Results, Sections 2–5; Tables 2–5 | Reports factor structure, model fit, reliability, GDHRI-outcome associations, confidence intervals, p-values, adjusted $R^2$, and incremental explanatory power. |
| 17 | Other analyses | Results, Section 5; Table 5 | Reports fixed-effects panel models, random forest validation, equal-weighting sensitivity analysis, exclusion of high-income countries, and complete-case comparison. |
| 18–21 | Key results, limitations, interpretation, and generalisability | Discussion | Summarizes findings, interprets ecological associations cautiously, discusses limitations, and addresses external validity. |
| 22 | Funding | Funding statement | States that no external funding was received. |

† STROBE = Strengthening the Reporting of Observational Studies in Epidemiology. This checklist was adapted for a cross-national ecological study using publicly available aggregate country-level data. Items not applicable to individual-level cohort or case-control designs were omitted or addressed as not applicable where relevant.

**ORIGINAL ARTICLE**
Journal of Science Technology Education Art and Medicine
DOI: 10.63137/jsteam.241906

**Table S2: Digital Behavioral Search Terms and Wikipedia Page Titles Used for Data Retrieval (2015–2022)**

**Panel A: Google Trends Keywords**

| Category | Keyword | Retrieval Tool | Time Period | Geographic Scope |
|---|---|---|---|---|
| Symptoms | fever, cough, shortness of breath, fatigue, sore throat | gtrendsR (R package) | Jan 2015–Dec 2022 | Country-level RSV |
| Prevention | mask, handwashing, sanitizer, social distancing, quarantine | gtrendsR (R package) | Jan 2015–Dec 2022 | Country-level RSV |
| Healthcare services | hospital, doctor, ICU, testing | gtrendsR (R package) | Jan 2015–Dec 2022 | Country-level RSV |
| Vaccination | vaccine, vaccination, booster dose | gtrendsR (R package) | Jan 2015–Dec 2022 | Country-level RSV |
| Information-seeking | COVID symptoms, vaccine side effects, pandemic | gtrendsR (R package) | Jan 2015–Dec 2022 | Country-level RSV |

**Panel B: Wikipedia Page Titles**

| Category | Page Title | Retrieval Tool | Time Period | Aggregation |
|---|---|---|---|---|
| Vaccination | Vaccine, Vaccination | Wikimedia REST API | Jan 2015–Dec 2022 | Geolocated pageviews |
| Disease | Infection, COVID-19, Coronavirus disease | Wikimedia REST API | Jan 2015–Dec 2022 | Geolocated pageviews |
| Public health | Public health, Pandemic | Wikimedia REST API | Jan 2015–Dec 2022 | Geolocated pageviews |
| Prevention | Quarantine, Face mask, Hand washing | Wikimedia REST API | Jan 2015–Dec 2022 | Geolocated pageviews |

† RSV = relative search volume. Google Trends keywords were queried in English; for non-English-speaking countries, Google Trends returns results based on the dominant language of the country's Google domain. Wikipedia page titles correspond to English-language Wikipedia articles; pageview data were aggregated at the country level using geolocated access records where available. All terms were specified a priori.

ORIGINAL ARTICLE
Journal of Science Technology Education Art and Medicine
DOI: 10.63137/jsteam.241906



**Table S3. Composition and Weighting of the Global Digital Health Resilience Index (GDHRI)**

| Domain | Indicator | Data Source | Baseline Period | PCA Weight | Scoring Direction |
|---|---|---|---|---|---|
| Digital Engagement Capacity | Internet penetration (%) | International Telecommunication Union (ITU) | 2019 | 0.18 | Higher = better |
| Digital Engagement Capacity | Digital Adoption Index | World Bank Open Data | 2019 | 0.21 | Higher = better |
| Digital Engagement Capacity | Google Health RSV | Google Trends API (gtrendsR) | 2015–2019 average | 0.16 | Higher = better |
| Information Ecosystem Integrity | Media freedom index | Reporters Without Borders (RSF) | 2019 | 0.13 | Higher = better |
| Information Ecosystem Integrity | Trust in information sources | World Values Survey Wave 7 (Variable V158: Confidence in the Press) | 2017–2022 nearest available baseline | 0.12 | Higher = better |
| Structural Health System Capacity | Hospital beds per 1,000 | World Bank World Development Indicators (WDI) | 2019 | 0.07 | Higher = better |
| Structural Health System Capacity | Physicians per 1,000 | World Bank World Development Indicators (WDI) | 2019 | 0.08 | Higher = better |
| Structural Health System Capacity | UHC service coverage index | WHO Global Health Observatory (GHO) | 2019 | 0.05 | Higher = better |

† GDHRI = $\Sigma(w_i \times z_i)$, rescaled to 0–100. Weights derived from principal component analysis with varimax rotation. RSV = relative search volume; ITU = International Telecommunication Union; RSF = Reporters Without Borders; WDI = World Development Indicators; GHO = Global Health Observatory; UHC = Universal Health Coverage.