**Supplementary Methods**


**Authors:**

Xinyu Zhou,[a] Xu Zhang,[a] Heidi J. Larson,[b] Alexandre de Figueiredo,[b] Mark Jit,[b] Samah Fodeh,[c] Sten H Vermund,[d] Shujie Zang,[a] Leesa Lin,[b] Zhiyuan Hou,[a]

**Affiliations:**

[a] School of Public Health, NHC Key Laboratory of Health Technology Assessment, and Global Health Institute, Fudan University, 130 Dong'an Road, Shanghai, 200032, China.

[b] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, England.

[c] Department of Emergency Medicine, Yale School of Medicine, New Haven, USA.

[d] Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, USA.

**Correspondence to** Zhiyuan Hou (email: zyhou@fudan.edu.cn).

# Supplementary Methods

**Table of contents:**

## Supplementary Methods

This section provides a detailed description of our methods. To leverage deep learning models for analyzing multilingual X posts, we first developed an annotation framework for vaccine sentiment and had humans annotate a sample of X posts according to the annotation framework; then we finetuned multilingual deep learning models to imitate human annotation; and finally, using the finetuned deep learning models, we annotated all X posts.

### Multilingual X data collection

#### Identify multilingual keywords concerning the COVID-19 vaccine

To collect multilingual data on X concerning the COVID-19 vaccine, the first step is to identify all relevant keywords. In this step, we collected, expended, and verified keywords produced by professional translators from a large translation company in China (Beijing Chinese-Foreign Translation & Information Service Co., Ltd., "译鱼人工翻译", http://www.cipgtrans.com/).

Firstly, we instructed the company to identify translators for 97 languages (the 100 languages supported by the XLM-RoBERTa (XLM-R)(1) model except for three languages known by the authors: English, Traditional Chinese, and Simplified Chinese). The company identified translators in 87 languages but could not find experts for the remaining 10. We provided these translators with sample keywords and detailed instructions in both Chinese and English, directing them to list all relevant keywords associated with the COVID-19 vaccine in their respective languages. From this, they identified 483 keywords across the 87 languages. Separately, our research team collected 14 keywords in English, Simplified Chinese, and Traditional Chinese. In total, we gathered 497 keywords across 90 languages.

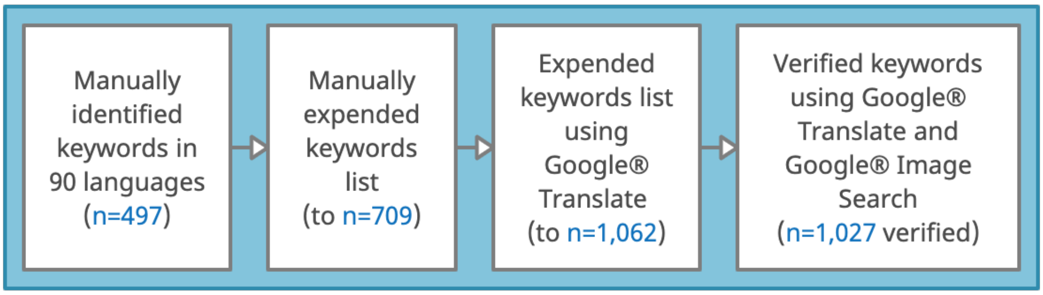
In numerous languages, the terms "COVID", "COVID-19", and "COVID19" are synonymous. Nevertheless, there is a possibility that the translators might recognize only one or two of these as a part of the keywords, overlooking the others. To address this, researchers have decided to review the original list of manually translated keywords. If they find any instance of "COVID-19", "COVID19", or "COVID", they will proceed to include any missing variations to the keyword list. Through this meticulous approach, we have successfully expanded our list of manually translated keywords, bringing the total to 709 keywords.

To enrich our keyword list, we utilized Google Translate to obtain machine translations of nine handpicked English keywords related to the COVID-19 vaccine. These keywords are “COVID vaccine”, “COVID vaccines”, “COVID vaccination”, “COVID-19 vaccine”, “COVID-19 vaccines”, “COVID-19 vaccination”, “coronavirus vaccine”, “coronavirus vaccines”, and “coronavirus vaccination”. We gathered machine translations for 86 languages, all of which had been translated by our translators and supported by Google Translate. The only exception was Burmese, which, as of February 2022, was not supported by Google Translate. Duplicated keywords from manual and machine translations were removed, and our keyword list was expanded to 1062 keywords.

In the final step, our team rigorously verified the relevance of all the collected keywords to the COVID-19 vaccine. We employed Google Translate to revert all keywords from the 87 languages to English and Chinese. Any keyword not translated to "COVID-19 vaccine/vaccination" or equivalent in either language was discarded. Additionally, we utilized Google Image Search to authenticate the retained keywords. Those that displayed more than two unrelated images to the COVID-19 vaccine on the first page of search results (which contains around 20 images) were also omitted. Consequently, we confirmed the eligibility of keywords across 90 languages, amounting to a total of 1027 (as detailed in the Supplementary Worksheet).

A visual representation of the keyword identification strategy can be found in Figure S1.

**Figure S1.** Strategy for identifying multilingual keywords concerning the COVID-19 vaccine




**Collection of COVID-19 vaccine-related X posts by Meltwater platform**

Using the multilingual keywords verified above, we collected 13,093,406 multilingual posts on X in the public domain regarding COVID-19 vaccination across the globe from November 13, 2020, to March 5, 2022, through Meltwater media monitoring and social listening platform. Meltwater platform has contracts with various social media platforms and has permission to access X data.

During our data collection phase, we gathered original X posts as well as quotes. Conversely, re-posts, which don't permit user comments, were excluded from our collection. We also opted not to collect replies. The distinctions between reposts and quotes are elaborated upon in the X Help Centre. To put it succinctly, a repost involves sharing another user's X post in its original form (as-is), whereas a quote enables a user to append their own comment while sharing an X post. X posts with identical content from the same user were deemed duplicates and subsequently removed.

Meltwater leverages X users' profile location to provide the best knowledge of users' geo-location. We used X posts with geo-location data to assess global spatial variations in COVID-19 vaccination opinions expressed among X users across countries and territories. We focused on countries and territories meeting the following criteria: (1) X is not banned by the relevant government, (2) there are data on COVID-19 vaccine from at least 100 X users in that country or territory, (3) the analyzed 90 languages cover all official languages in that country or territory. There were 171 countries and territories with posts sent by at least 100 X users, and their censorship status and official languages are shown in the Supplementary Worksheet. After excluding countries or territories with official languages not covered by our dataset and those banning X, 135 countries or territories were included; when stratified by WHO region, they are 12 countries or territories from the Western Pacific Region (WPR), six from the Southeast Asian Region (SEAR), 41 from the European Region (EUR), 18 from the Eastern Mediterranean Region (EMR), 31 from the Americas Region (AMR), and 27 from the African Region (AFR).

### Annotation framework for vaccine acceptance

We used the confidence, complacency, and convenience (i.e. 3Cs) model(2) of vaccine hesitancy proposed by the World Health Organization (WHO) to develop an annotation framework for COVID-19 vaccine-related posts, which was validated in a sample of 500 posts. Vaccine acceptance and vaccine refusal were the core measures. In addition, we investigated determinants of vaccine acceptance, such as confidence in vaccines, the online information environment, and perceived barriers to accessing vaccines. Specifically, our annotation framework covered four key concepts related to COVID-19 vaccination and included eight categories (Table 1). First, COVID-19 vaccine acceptance covered the categories of: (i) intent to accept vaccination; and (ii) intent to refuse vaccination. Second, confidence in COVID-19 vaccines covered: (iii) belief that vaccines are effective; (iv) belief that vaccines are not safe; and (v) distrust in government. Third, the online information environment regarding COVID-19 vaccines covered: (vi) misinformation or rumours about vaccines. Fourth, perceived barriers to accessing COVID-19 vaccines covered: (vii) vaccine accessibility; and (viii) vaccine equity. The definition of each annotation category relating to COVID-19 vaccination is available in Table S1a.

**Table S1.** Annotation framework and examples

**Table S1a.** The annotation framework and its eight predefined categories for annotating COVID-19 vaccine-related X posts.

| Annotation category | Definition |
|---|---|
| Vaccine acceptance | |
| (a) Intent to accept COVID-19 vaccination | X posts indicating that they will accept, support or be willing to get COVID-19 vaccination. |
| (b) Intent to reject COVID-19 vaccination | X posts indicating that they will reject, do not support, or be unwilling to get COVID-19 vaccination. |
| Vaccine confidence | |
| (c) Belief that COVID-19 vaccine is effective | X users had confidence in the effectiveness of the COVID-19 vaccine, believing that it is effective. |
| (d) Belief that COVID-19 vaccine is not safe | X users lacked confidence in the safety of the COVID-19 vaccine, believing it was not safe. |
| (e) Distrust in government | X users indicated distrust in government or policymakers, including all-level government, ministry of health, CDC, etc. |
| Information environment | |
| (f) Misinformation or rumours on COVID-19 vaccine | Negative information about vaccines in X posts, such as misinformation, rumours, anti-vaccine campaigns, anti-intellectual, anti-science campaigns, and vaccine scandals. |
| Perceived barriers to accessing vaccines | |
| (g) COVID-19 vaccine accessibility | X posts discussed and concerned the production or supply capacity of the COVID-19 vaccine or the self-efficacy of accessing it. |
| (h) COVID-19 vaccine equity | X posts discussed and concerned (priority) vaccination groups or vaccine allocation equity. |

**Table S1b.** Examples of X posts and their annotation categories.

| X post | Annotation |
|---|---|
| **QT: #VaccinesWork #VaccinesSaveLives; The #MMIPConference highlights the critical role of the COVID-19 Vaccine Taskforce which has brought together industry & academic partners, working at pace to accelerate manufacture of leading vaccine candidates. We're very proud to be part of this valuable initiative!** | (a), (c) |
| **QT: Bill did say "if successful, along with birth control, it will reduce the population by 10-15%..." His words not mine #ExposeGates; What if the COVID-19 vaccine actually kills more people than COVID-19?** | (b), (d) |
| **QT: Russia & Trump are trying to make sure this virus doesn't go away smh!; BREAKING: A cyber espionage group, almost certainly part of the Russian intelligence services attempted to hack coronavirus vaccine research, U.K., U.S. and Canada allege.** | (e) |
| **QT: Vaccine inequality and lack of access will be our next global crisis; Important call from world leaders for equitable to a Covid-19 vaccine - this is a global crisis that can only be solved with a joined up, global response** | (a), (g), (h) |
| **QT: That's very white of Melinda and pasty old white bread #ExposeBillGates to experiment on minority communities first w/a bypass-animal-testing #vaccine like US Gov did on #Tuskegee Airmen. What will the #GatesFoundation think of next to thin the herd of humanity?; Melinda Gates Wants to Deliver Coronavirus Vaccine Based on Racial Groups, Blacks First** | (e), (f), (h) |
| **One issue around the potential COVID19 vaccine is that big pharma is more incentivised to develop a good, yet not GREAT vaccine, because requiring a booster shot or two every year would create massive recurring revenues. Say 1.5bn people x $39 every year, that's $58.5b / year.** | *This X posts doesn't belong to any category.* |

Note: (a-h) categories are shown in Table S1a.

## Manual annotation of the sampled English-language X posts

Using this framework, two annotators independently annotated 8125 English-language posts on COVID-19 vaccination. Any disagreement was resolved by a third annotator. There were two main steps: (i) each annotator separated human-generated posts from news reports, advertisements, government announcements and posts generated by automated (i.e. bot) accounts; and (ii) each human-generated post was annotated according to its relevance to the eight annotation framework categories. A post could be relevant to one or more categories or none. Examples of annotated posts are available in Table S1b.

## Finetuning multilingual deep learning models using manually annotated X posts

### About the XLM-R deep learning model

Bidirectional Encoder Representation from Transformer (BERT), proposed by Google in 2018(3), is a "large" pre-trained natural language processing model that outperformed all models before it on the General Language Understanding Evaluation benchmark. Meanwhile, its multilingual version – multilingual BERT, pre-trained using massive multilingual data, demonstrated unprecedented cross-lingual comprehension capacity. Remarkably, when fine-tuned with a modest dataset of manually annotated task-specific data in one language, multilingual BERT could proficiently handle the same task in approximately 100 languages without requiring translations software. Multilingual BERT offers a new possibility for analyzing cross-lingual data, and since then, numerous multilingual models have been developed.

To extend our manual classifications, we applied XLM-RoBERTa (XLM-R), a state-of-the-art multilingual model that outperformed multilingual BERT significantly, especially in low-resource languages(1). XLM-R is a cutting-edge pretrained transformers-based deep neural network for analyzing multilingual textual data. In diversified benchmarks for multilingual models, it significantly outperformed other widely-used state-of-the-art multilingual models, such as XLM and multilingual BERT, especially in low-resource languages, such as Urdu and Swahili(1). In these benchmarks, multilingual models' performance was evaluated in multi-lingual datasets when finetuned purely with mono-lingual datasets. The XLM-R model was pre-trained using the Masked Language Models objective in a self-supervised style, using 2.5 TB of filtered CommonCrawl data in the 100 most spoken languages. During the pre-training process of the model, language materials in each of the 100 languages were provided to the model, and some parts (15%) of the text were randomly masked. The model is asked to predict the masked tokens as correctly as possible, and after intensive pretraining, it comprehends 100 languages simultaneously. As a result, once the pre-trained model is fine-tuned with a small, manually annotated monolingual downstream training data, it can precisely analyze multilingual data in around 100 languages. In the Cross-lingual Natural Language Inference test, the XLM-R model finetuned with English-language data reached 80.9% accuracy in cross-lingual transfer, outperforming XLM by 10.2% and multilingual BERT by 14.6%; it not only reached high accuracy in high-resource languages (English and French at 89.1% and 84.1%, respectively) but also performed well in low-resource languages (Swahili and Urdu at 73.9% and 73.8%, respectively). In this study, we employed and finetuned the 24-layer version of the pre-trained XLM-R model, which contains 550 million parameters.

**Data augmentation and hyperparameters for XLM-R deep learning model**

To deploy the model for analyzing X posts surrounding the COVID-19 vaccine, the pre-trained XLM-R model needed to be finetuned using our manually annotated COVID-19 X posts dataset. We randomly selected around 80% of our annotated X posts as the training set, around 10% as the validation, and around 10% as the test set. Given the imbalance in our manually annotated data, we utilized data augmentation to increase the count of positive labels in the training set. This step was crucial to bolster the model's performance prior to its fine-tuning. We hired two approaches for data augmentation: 1) back translation (translating a text to another language, then back to the original) using Baidu Translate API (http://api.fanyi.baidu.com/), and 2) simulating spelling errors using the python package “nlpaug”. We chose the models' hyperparameters based on the model’s performance in our validation set. The hyperparameters are available in Table S2.

**Table S2.** Settings (hyperparameters) to fine-tune the XLM-R deep learning models.

| XLM-RoBERTa Model (a): find human-generated posts, N(training set)=**6581**, N(validation set)=**732** | | | | | | |
|---|---|---|---|---|---|---|
| Annotation category | n(labelled) before data augmentation | n(labelled) after data augmentation | batch size | learning rate | epoch | Data augmentation |
| Human-generated | 3753 | 3753 | 16 | 2.00E-05 | 4 | no |
| XLM-RoBERTa Model (b): classify, N(original training set)=**3377**, N(validation set)=**376** | | | | | | |
| Annotation category | n(labelled) before data augmentation | n(labelled) after data augmentation | batch size | learning rate | epoch | Data augmentation |
| Vaccine acceptance | | | | | | |
| (a) Intent to accept COVID-19 vaccination | 1911 (1720+191) | 1911 (1720+191) | 32 | 2.00E-05 | 4 | no |
| (b) Intent to reject COVID-19 vaccination | 489 (440+49) | 1369 (1320+49) | 24 | 2.00E-05 | 3 | yes |
| Vaccine confidence | | | | | | |
| (c) Belief that COVID-19 vaccine is effective | 671 (604+67) | 1275 (1208+67) | 16 | 1.00E-05 | 4 | yes |
| (d) Belief that COVID-19 vaccine is not safe | 300 (270+30) | 570 (270+30) | 32 | 2.00E-05 | 3 | yes |
| (e) Distrust in government | 445 (400+45) | 1245 (1200+45) | 16 | 2.00E-05 | 4 | yes |
| Information environment | | | | | | |
| (f) Misinformation or rumours on COVID-19 vaccine | 345 (310+35) | 1275 (1240+35) | 32 | 2.00E-05 | 3 | yes |
| Perceived barriers to accessing vaccines | | | | | | |
| (g) COVID-19 vaccine accessibility | 297 (267+30) | 831 (801+30) | 16 | 2.50E-05 | 3 | yes |
| (h) COVID-19 vaccine equity | 587 (528+59) | 1115 (1056+59) | 24 | 2.00E-05 | 4 | yes |

**Measuring the performance of XLM-R deep learning models: $F_1 - score$ and $precision$**

$F_1 - score$ and $precision$ are widely used metrics in Natural Language Processing to measure the performance of machine learning models. The XLM-R models were asked to predict the human annotations in the held-out test set, and we compared the machine's annotation with the human's annotation. Here, human annotation was considered the ground truth, a common practice in machine learning. Like other studies, $F_1 - score$ and $precision$ of our models are calculated as follows.

$$F_1 - score = 2 \cdot \frac{1}{\frac{1}{precision} + \frac{1}{recall}}$$

Where,

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

In the test set, we evaluated the performance of XLM-R models. As shown in Table S3, the $precision$ of the models used in this study ranged from 61.76% to 89.74%, whereas the $F_1 - scores$ ranged from 67.53% to 86.62%.

**Table S3.** Performances of XLM-R deep learning models in predicting each of eight predefined categories relating to COVID-19 vaccination.

- 95% Confidence Intervals (CI) were estimated using the law of large number, i.e.

$$\hat{p} \pm Z_\alpha \cdot \sqrt{\frac{\hat{p}\,(1 - \hat{p})}{n}}$$

($Z_\alpha = 1.96$ for the error rate $\alpha = 0.05$)

| **Finetuned XLM-RoBERTa Model (a): to find human-generated posts, N(test set)=812** | | | |
|---|---|---|---|
| Annotation category | n | $F_1$-score (and 95% CI) | Precision (and 95% CI) |
| Human-generated | 435 | 0.8662<br>[0.8342, 0.8982] | 0.8782<br>[0.8474, 0.9089] |
| **Finetuned XLM-RoBERTa Model (b): to classify, N(test set)=435** | | | |
| Annotation category | n | $F_1$-score (and 95% CI) | Precision (and 95% CI) |
| Vaccine acceptance | | | |
| (a) Intent to accept COVID-19 vaccination | 222 | 0.8541<br>[0.8076, 0.9005] | 0.8964<br>[0.8563, 0.9365] |

| | | | |
|---|---|---|---|
| (b) Intent to reject COVID-19 vaccination | 54 | 0.7304<br>[0.6121, 0.8488] | 0.7778<br>[0.6669, 0.8887] |
| Vaccine confidence | | | |
| (c) Belief that COVID-19 vaccine is effective | 82 | 0.8098<br>[0.7249, 0.8948] | 0.8049<br>[0.7191, 0.8907] |
| (d) Belief that COVID-19 vaccine is not safe | 34 | 0.6753<br>[0.5179, 0.8327] | 0.7647<br>[0.6221, 0.9073] |
| (e) Distrust in government | 55 | 0.7921<br>[0.6848, 0.8993] | 0.7273<br>[0.6096, 0.8450] |
| Information environment | | | |
| (f) Misinformation or rumours on COVID-19 vaccine | 34 | 0.7500<br>[0.6044, 0.8956] | 0.6176<br>[0.4543, 0.7810] |
| Perceived barriers to accessing vaccines | | | |
| (g) COVID-19 vaccine accessibility | 41 | 0.6818<br>[0.5392, 0.8244] | 0.7317<br>[0.5961, 0.8673] |
| (h) COVID-19 vaccine equity | 64 | 0.8088<br>[0.7125, 0.9052] | 0.8594<br>[0.7742, 0.9445] |

### Deep learning-based annotation of all multilingual X posts

Finally, we used our fine-tuned XLM-R model to annotate all COVID-19 vaccine-related X posts (n=13,093,406) automatically. Like our manual annotation process, the models first identified if each X post was sent by humans or not. The X posts deemed not to be sent by an individual human (n=6,046,183) were excluded from further analysis. The remaining human-generated X posts (n=7,047,223) were further classified (annotated) using fine-tuned XLM-R models.

## External data utilized in this study

**Table S4.** Country-level indicators used in regression analysis on the determinants of COVID-19 vaccination acceptance and coverage across countries.

| Indicators | Description | Data source | Time range | Value range |
|---|---|---|---|---|
| **COVID-19 vaccination coverage** | | | | |
| **COVID-19 vaccination coverage, %** | The total number of doses divided by the total population of the country. | Our World in Data | Nov, 2020 – Mar, 2022 | .48-121.53 |
| **Governance** | | | | |
| **Government effectiveness** | Perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. | World Bank | 2020 | -2.31-2.34 |
| **Control of corruption** | Perceptions of the extent to which public power is exercised for private gain, including petty and grand forms of corruption, and "capture" of the state by elites and personal interests. | World Bank | 2020 | -1.71-2.27 |
| **State fragility** | Indicator P2: Incapacity to provide essential public goods and services and cope with shocks | Fragility States Index | 2022 | .90-10 |
| **Pandemic preparedness** | | | | |
| **Epidemic Ready Score** | The ready score determines whether a country is prepared to find, stop, and prevent epidemics using the WHO's Joint External Evaluation data. There are 19 areas of preparedness and response capacity scored. | Prevent Epidemics | Latest | 26-93 |
| **Global Health Security Index** | The GHS index includes six categories: prevention, detection and reporting, rapid response, health systems, compliance with international norms, and risk environment. Although the GHS Index can identify preparedness resources and capacities available in a country, it cannot predict whether or how well a country will use them in a crisis. | GHS Index | 2021 | 16-75.90 |

| | | | | |
|---|---|---|---|---|
| **Doctors per 1000 people** | Health systems resources | The Global Health Observatory indicators | Latest | .23-84.20 |
| **Trust** | | | | |
| **Trust in government, %** | Trust coded as “A lot” or “Some” on W5B asking about confidence in government | Wellcome Global Monitor 2020 COVID-19 | 2017-2021 | 16.14-93.70 |
| **Trust in science, %** | Trust coded as “A lot” or “Some” on W6 asking about trust in science | Wellcome Global Monitor 2020 COVID-19 | 2018 | 47.17-96.70 |
| **Interpersonal trust, %** | Trust coded as “most people can be trusted” on Q57 asking about interpersonal trust | World Values Survey Wave 7 | 2021 | 4.25-56.58 |
| **Culture-related index** | | | | |
| **Individualism** | The ties between individuals are loose: everyone is expected to look after him/herself and his/her immediate family | Clearly Cultural | Latest | 6-91 |
| **Uncertainty avoidance** | Uncertainty avoidance deals with a society’s tolerance for uncertainty and ambiguity. It indicates to what extent a culture programs its members to feel either uncomfortable or comfortable in unstructured situations (novel, unknown, surprising, and different from usual). | Clearly Cultural | Latest | 8-112 |
| **Social development status** | | | | |
| **Socio-Demographic Index** | The Socio-demographic Index (SDI) is a composite indicator of development status strongly correlated with health outcomes. It is the geometric mean of 0 to 1 indices of total fertility rate under the age of 25 (TFU25), mean education for those ages 15 and older (EDU15+), and lag distributed income (LDI) per capita. | Global Burden of Disease | 2019 | .08-0.93 |

| | | | | |
|---|---|---|---|---|
| **Ln (GDP per capita)** | Natural logarithm of GDP per capita (constant US＄) | World Bank | 2021 | 6.10-11.82 |
| **School enrolment, %** | The ratio of total enrollment, regardless of age, to the population of the age group. | World Bank | 2020 | 1-149 |
| **Internet coverage, %** | Proportion of individuals using the Internet in total population. The Internet can be used via a computer, mobile phone, personal digital assistant, games machine, digital TV, etc. | World Bank | 2020 | 10-100 |
| **Demographics** | | | | |
| **Population density** | People per sq. km of land area | World Bank | 2020 | 2-8019 |
| **Population ages 0-14, %** | Percentage of population aged 0-14 in total population | World Bank | 2021 | 12.27-46.73 |
| **Population ages 65 and above, %** | Percentage of population aged 65 and above in total population | World Bank | 2021 | 1.45-28.70 |
| **Urban population, %** | Percentage of urban population in total population | World Bank | 2021 | 18.86-100 |

**Table S5.** Country policy indicators used in panel data analysis on the determinants of temporal trends in COVID-19 vaccine acceptance across countries.

| Indicators | Description | Data source | Value range |
|---|---|---|---|
| **AEFI report** | Serious adverse events following COVID-19 vaccination have been reported. | News websites and official government websites | 0, 1 |
| **Vaccine availability** | Availability of COVID-19 vaccines to the public in the country during the week. | The Oxford Covid-19 Government Response Tracker | 0, 1 |
| **Mandatory vaccination** | Existence of COVID-19 vaccination requirements for one or more groups | The Oxford Covid-19 Government Response Tracker | 0, 1 |
| **Adolescent vaccination eligibility** | Expand COVID-19 vaccination eligibility for adolescents aged 5-15 years | The Oxford Covid-19 Government Response Tracker | 0, 1 |
| **Sufficient vaccine supply** | Whether the supply of COVID-19 vaccines surpasses the country's population | Our World in Data; World Population Prospects | 0, 1 |
| **Relaxation of NPI measures** | Adoption of relaxed Non-Pharmaceutical Intervention (NPI) measures in the country. | News websites and official government websites | 0, 1 |

1. Conneau A, Khandelwal K, Goyal N, Chaudhary V, Wenzek G, Guzmán F, et al. Unsupervised cross-lingual representation learning at scale. arXiv preprint arXiv:191102116. 2019.
2. MacDonald NE. Vaccine hesitancy: Definition, scope and determinants. Vaccine. 2015;33(34):4161-4. Epub 2015/04/22.
3. Devlin J, Chang M-W, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:181004805. 2018.