THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Causey K, Fullman N, Sorensen RJD, et al. Estimating global and regional disruptions to routine childhood vaccine coverage during the COVID-19 pandemic in 2020: a modelling study. *Lancet* 2021; published online July 15. http://dx.doi.org/10.1016/S0140-6736(21)01337-4.

Supplementary appendix to Estimating global and regional disruptions to routine childhood vaccine coverage during the COVID-19 pandemic in 2020: a modelling study

This appendix provides more detailed methodology and supplemental figures, tables, and results for "*Estimating global and regional disruptions to routine childhood vaccine coverage during the COVID-19 pandemic in 2020: a modelling study.*" Portions of this appendix have been adapted from Galles and colleagues,¹ and as well as Zheng and colleagues.² References are provided for adapted sections.

Table of contents

Section 1: Compliance with Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER)

Section 2: Data processing

Section 2.1: Data inclusions and exclusions

Supplementary table 1 details all sources included in the analysis, including data from the Asociación Española de Vacunología,³ DHIS-2 in Nepal,⁴ HMIS in India,⁵ NCIRS in Australia,⁶ NHMIS in Nigeria,⁷ Public Health England,⁸ Public Health Scotland,⁹ Scientific American,¹⁰ Virginia Mercury,¹¹ and data reported by countries to WHO Regional Offices.12 We included data for the first dose of measlescontaining-vaccine (MCV1) and third dose of diphtheria-tetanus-pertussis (DTP3) for 2020, using 2019 and/or January-February 2020 as the reference period.

Data were found through internet searches and professional contacts. We also reviewed all of the studies in a systematic review¹³ and a curated list of relevant publications¹⁴ for data that met inclusion criteria.

For the data that national governments reported to WHO regional offices, we calculated monthly cumulative disruption ratios (CDRs) based on the monthly number of doses administered in 2019-2020. For several countries in Africa ($N = 22$) we were missing monthly doses delivered for September-December of 2019. We imputed values for these months by subtracting the sum of all available monthly values of doses delivered (January – August) from the total annual 2019 doses delivered as reported to the WHO via the Joint Reporting Form¹⁵ and dividing by 4 to create a monthly average estimate number of doses delivered in the last quarter of 2019.

After imputation, we calculated 1402 CDRs. We excluded CDRs that did not align with qualitative sources of disruption data,^{16,17,18} likely included bias in the final month due to reporting delays, or countries for which we had an alternative source of administrative data. Exclusions are detailed in the figure below:

We decided to incorporate the adjustment for reporting delays after seeing evidence of lagged reporting across sequential releases of administrative data. In several countries, we saw evidence of under-reporting in the last month of available data, suggested by low numbers of doses delivered in an early data release that were adjusted upwards in a subsequent data release. As a result, we decided to exclude any CDRs from the terminal month of the time series where we saw a decrease of greater than 20% between the 2020 to 2019 ratio calculated for the second-to-last month of data compared to the ratio calculated for the last month of data.

Additionally, we conducted a sensitivity analysis, excluding all last months of data from the administrative data reported to WHO Regional offices to understand the potential impact on our results. Results from this analysis are provided in section 4.5.

The following table details the data included by Global Burden of Disease (GBD) super-region:

Section 2.2: Imputation for missing vaccines

In location-months missing one of the two vaccines, due to data exclusions or missingness, we calculated CDRs based on the available vaccine and the ratio of CDRs for DTP3 and MCV1.

For Antigua and Barbuda, Costa Rica, Equatorial Guinea, Indonesia, Madagascar, Saint Lucia, and Uganda, we had overlapping data of both vaccines to calculate location-specific ratios, MCV1 over DTP3, of 1.27, 0.57, 1.16, 1.03, 1.13, 0.92, and 1.03 respectively. To impute for Alabama (USA), Central African Republic, Kenya, Lesotho, Namibia, Senegal, Somalia, and Tanzania, we calculated a global ratio based on 608 pairs in 90 locations; the average disruption ratio was 1.01 for MCV1 to DTP3.

Section 2.3: Data variance calculations

For each data point, we assigned a weight in the model equal to the inverse variance of the log ratio.¹⁹ For sources that provided monthly doses in 2019 and 2020 (DHIS-2 in Nepal,⁴ HMIS in India,⁵ NHMIS in Nigeria,⁷ Public Health England,⁸ Public Health Scotland,⁹ and WHO Regional Offices¹²), we calculated the variance of the log cumulative ratio for month, m , with the following formula:²⁰

By using this formula, the variance decreased over time as the number of doses increased. While this method provides a quantitative estimate of the variance for each administrative data point, it is difficult to fully quantify all potential sources of uncertainty in administrative data systems (i.e. due to variation in the reliability of administrative reporting systems). For instance, estimates calculated based on administrative data in large population countries had implausibly small variance due to the high number of vaccines administered. We therefore set a monthly minimum and maximum on the standard error of the log ratio, equal to the $20th$ and $80th$ percentiles of all available data by month. We used September quantities for October-December, only calculating values for months with at least 100 CDRs. This minimum affected the standard error of all 20 CDRs from the NHMIS data,⁷ 256 CDRs from the WHO Regional Office data,¹² and all 8 CDRs from the India HMIS data,⁵ while the maximum affected 238 CDRs from the WHO administrative data.¹²

For estimates without sample size or sufficient information to calculate the variance, we imputed the median of the standard error of the log cumulative ratio by month for sources based on administrative data, electronic medical records, and immunization registries ($N = 18$, sources: Asociación Española de Vacunología³ and NCIRS in Australia⁶) and the 80th percentile of the standard error of the log CDR by month for sources published in news media ($N = 58$, sources: Scientific American¹⁰ and Virginia M ercury¹¹). We used September quantities for October-December.

Section 2.4: Subnational data

In Spain, the United Kingdom, and the United States of America, there were not sources of nationallyrepresentative data, but we did have access to data from some administrative regions in Spain, Scotland and England of the UK, and several US states (supplementary table 1). Because we had mobility disruption data aligned to these subnational locations, we included all subnational-level data in the step 1 national model for the corresponding country. For example, paired observations of vaccine coverage and mobility disruptions from US states were included in the step 1 model for the US to inform the relationship between vaccine CDRs and cumulative mobility disruptions (figure 2).

In the step 2 model of residuals, we only included data at the national level. In the US and the UK, the locations included represented over 50% of the national population; we population-weighted the available subnational CDRs to calculate approximate national-level CDRs for each month. The residuals were calculated at the national level as inputs to the step 2 model. For Spain, the available data did not provide sufficient coverage of the national population, so we elected to use the location-specific step 1 model and the super-region step 2 model to make estimates in Spain

Section 2.5: Data coverage

The following table enumerates the data coverage for both routine immunisation (RI) disruptions (rows) and mobility data (columns). It displays the number of countries and the fraction of the target population that fall into each data-availability category. We have mobility data in 69% of countries (*n*=134), accounting for 87% of the target population for both DTP3 and MCV1. The majority of children in the target populations for each vaccine live in countries where we have data on both changes in human mobility and RI disruptions. Less than 4% of children in the target population live in countries where we have data on neither.

Section 3: Spline cascade for continuous data

Section 3.1: Spline modeling²

In this section we discuss spline models for dose-response relationships. For general background on splines and spline regression see various sources. $21,22$

A spline basis is a set of piecewise polynomial functions with designated degree and domain. If we denote polynomial order by p, and the number of knots by k, we need $p + k$ basis elements s_j^p , which can be generated recursively as illustrated in the figure below.

Recursive generation of bspline basis elements (orders 0, 1, 2).

Given such a basis, we can represent any dose-response relationship as the linear combination of the spline basis elements, with coefficients $\boldsymbol{\beta} \in \mathbb{R}^{p+k}$:

$$
f(t) = \sum_{j=1}^{p+k} \beta_j^p s_j^p(t)
$$
 (1)

These coefficients are then inferred by the cascade model as discussed in section 3.4. An explicit representation of (1) is obtained by building a design matrix **X**. Given a set of t values at which we have data, the *j*th column of \bf{X} is given by the expression

$$
X_{\cdot}, j = \begin{bmatrix} s_j^p(t_0) \\ \vdots \\ s_j^p(t_k) \end{bmatrix} \tag{2}
$$

The model for direct observations of data coming from (1) can now be written compactly as

$$
y = X\beta + \epsilon
$$

Section 3.2: Linear mixed effects model with constraints and priors We consider the following linear mixed effects model:

$$
y_i = X_i(\beta) + Z_i u_i + \epsilon_i
$$

$$
u_i \sim N(0, \Gamma), \qquad \Gamma = diag(\gamma), \qquad \epsilon_i \sim N(0, \Lambda)
$$
 (3)

where $y_i \in R^{n_i}$ is the vector of observations from the *i*th study or source, $\epsilon_i \in R^{n_i}$ are measurement errors with given covariance Λ , $u_i \in R^{k_y}$ are independent random effects, and $Z_i \in R^{n_i \times k_y}$ is a linear map, and β are regression coefficients.

The linear mixed effects model can be extended to incorporate nonlinear inequality constraints

$$
C\theta\leq c,
$$

where $\theta = (\beta, \gamma)$.

We also allow priors on parameters of interest. We assume that priors are given by a functional form

$$
\theta \sim \exp(-\rho(\theta))
$$

The likelihood problem is then augmented and fit as described elsewhere.²

Section 3.3: Spline constraints

Using the constraints framework, we can impose shape constraints such as monotonicity, concavity, and convexity on splines.

Monotonicity. Spline monotonicity across the domain of interest follows from monotonicity of the spline coefficients.²² Given coefficients β , the curve $f(t)$ in (2) is *monotonically non-decreasing* when

$$
\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n
$$

and *monotonically non-increasing* if

$$
\beta_1 \ge \beta_2 \ge \cdots \ge \beta_n
$$

These are all linear inequality constraints, so easily incorporated into the framework of previously defined formulations for trimmed constrained mixed effects models.2

Convexity and concavity. For any twice continuously differentiable function $f : R \to R$, concavity and convexity are captured by the signs of the second derivative. Specifically, f is convex if $f''(t) > 0$ everywhere, and concave if $f''(t) \leq 0$ everywhere. We can compute $f''(t)$ for each interval, and impose linear inequality constraints on these expressions.

Rate constraints using derivatives. When fitting splines, it is useful to impose the constraint that pointwise rates of change are below plausible limits. The derivative of a spline is a linear function of the spline coefficients. ²² Therefore such a constraint is simply written as

$$
X_p \beta \leq L
$$

where X_p depends on the point in the domain as explained by (2), and L is the biologically plausible limit.

Enforcing linear tails. For large consumption with little data, we need the capacity to ensure that the last segment of the spline is linear, with slopes that match the adjacent segment at the knot. The estimated spline is then a best fit to the data, subject to this specification. Priors on the tails can also be provided.

Section 3.4: Cascade concept and notation

We develop notation and technical specification for a general cascade strategy useful for hierarchical models. We first set up a notation to explain the structure of the estimation.

Groups and levels. We are given a dataset consisting of tuples (y, x) , levels $\ell \in \{1, 2, \dots L\}$, and groups $i = 1, \ldots, N$. For example, in the figure below, there are three levels and nine groups. In global health, level 1 may represent global data, level 2 may correspond to the super-region, and so on.

Example of a 3-level cascade with 9 groups

Every datapoint is associated to k groups, one at each level of the hierarchy. For example, in the figure above, all datapoints corresponding to group 8 at level 3 also belong to group 4 at level 2 and to group 1.

We let I_{ℓ} enumerate groups *i* at level ℓ , with $n_{\ell} = |I_{\ell}|$, and

$$
N = \sum_\ell n_\ell
$$

For the example in the figure above, we have $I_1 = \{1\}$, $I_2 = \{2,3,4\}$, and $I_3 = \{5,6,7,8,9\}$, so that $n_1 = 1$, $n_2 = 3$, and $n_3 = 5$.

Every group i has a unique ancestor except for the root node, whose ancestor is empty. We associate each group to its ancestor using operator ↑. For the example in the figure above, we have

$$
1_{\uparrow} = \{\}
$$

$$
2_{\uparrow} = 3_{\uparrow} = 4_{\uparrow} = \{1\}
$$

$$
5_{\uparrow} = 6_{\uparrow} = \{2\}, 7_{\uparrow} = \{3\}, 8_{\uparrow} = 9_{\uparrow} = \{4\}
$$

Modeling cascade. For every level ℓ we fit n_l models, where for each $i \in I_\ell$ we consider data associated with that group. We indicate the level ℓ using superscript notation.

$$
y_i^{\ell} = X_i^{\ell} \beta_i^{\ell} + \epsilon_i^{\ell}, \qquad \epsilon_i^{\ell} \sim N(0, \Sigma_i^{\ell})
$$
\n
$$
\tag{4}
$$

We also allow level-specific priors on β_i^{ℓ} , which are informed by estimates from the parents:

$$
\beta_i^{\ell} = \hat{\beta}_{i_{\uparrow}}^{\ell-1} + \nu_i^{\ell}, \quad \nu_i^{\ell} \sim N(0, \lambda^{\ell} V_{i_{\uparrow}}^{\ell-1})
$$

For the special case of $\ell = 1$, we allow priors to be specified by the user. We also constrain β at all levels to lie in a set Ω , specified by equality and inequality constraints.

Translating to statements using negative log-likelihoods, for each group i we fit the model

$$
\hat{\beta}_i^{\ell} = \arg\min_{\beta \in \Omega} g(\beta; y_i^{\ell}, \Sigma_i^{\ell}) + p(\beta; \hat{\beta}_{i_1}^{\ell-1}, \lambda^{\ell} V_{i_1}^{\ell-1})
$$

In words, every group's fit is informed by data associated to that group along with priors whose mean and variance are obtained from those of the parent. The prior is allowed to depend on the level as well as on the parent, so that level-specific tuning parameters can be incorporated.

Gaussian example. The most common context is that of the Gaussian model (4), where we have

$$
\hat{\beta}_i^{\ell} = \arg \min_{\beta \in \Omega} \frac{1}{2} (y_i^{\ell} - X_i^{\ell} \beta)^T (\Sigma_i^{\ell})^{-1} (y_i^{\ell} - X_i^{\ell} \beta) + \frac{1}{2\lambda^{\ell}} (\beta - \hat{\beta}_{i_1}^{\ell-1})^T (V_{i_1}^{\ell-1})^{-1} (\beta - \hat{\beta}_{i_1}^{\ell-1})
$$
 (5)

The design matrices X_i^{ℓ} may represent linear covariates as well as splines, which are explained in section 3.1.

Section 3.5: Uncertainty quantification

Approximate posterior uncertainty intervals can be computed for any level of the cascade using asymptotic statistics. This approach is computationally efficient but does not incorporate constraints used to define the spline.

For the key example (5), the posterior variance estimate for β_i^{ℓ} is given by

$$
Var(\hat{\beta}_i^{\ell}) = \left(X_i^{\ell} (\Sigma_i^{\ell})^{-1} X_i + \frac{1}{2\lambda^{\ell}} (V_{i_1}^{\ell-1})^{-1} \right)^{-1}
$$

The approximate asymptotic uncertainty is used to inform more detailed uncertainty quantification procedures as described in section 4.2.

Section 4: Model details

Section 4.1: Modelling equations

We modeled the relationship between human mobility and the decrease in DTP3 and MCV1 vaccine coverage with two steps of a cascading random spline model with the Meta-Regression Bayesian, Regularized, Trimmed tool (MR-BRT).2 In step 1 we used the following set of equations to estimate the average relationship between the cumulative disruption ratio, *CDR*, and the cumulative mobility, *mob*:

Step 1: Disruption versus Mobility

Where *m* is the month, *v* is the vaccine – DTP3 or MCV1, *r* is the super-region, and *c* is the country. In stage 1 we fit a global model to all available ratios. We fit a cubic spline with a linear tail on the right side to the natural log of the ratio. We did not include an intercept, forcing the spline to intersect 1 (no disruption in vaccine coverage) at a mobility value of 0 (no change from baseline mobility). We also implemented a monotonicity prior, reflecting our prior assumption that vaccine coverage decreases as mobility decreases.

Each stage fit the same model to a subset of data, leveraging the estimates from the previous stage as a prior on the subsequent stage. In this way, the global model informed the super-region models; the superregion model informed the super-region-vaccine-specific model; and the super-region-vaccine-specific model informed the location-vaccine-specific model.

In the equations above, items in bold (β) represent matrices, and I is the identity matrix. Other values including (CDR, mob, and σ) are vectors, and N () represents the Gaussian distribution. The notation spline() refers to the design matrix described in section 3.2 equation 2.

While the step 1 model estimates the average relationship throughout the pandemic, it fails to account for the dynamic country-response over time. In step 2, we estimated the deviation from the average relationship over time by calculating the residuals between the data and model predictions over time. We used the following set of equations to estimate the residual, *r*, by month, *m*:

Where *m* is the month, *r* is the residual in logit space, and *p* is the prediction from the step 1 stage 4 results. Again fitting separate models by vaccine, we fit a global model to all available residuals, using a quadratic spline with linear tails. We included an intercept, β_0 , in all three stages and a location random effect, u_c , in stages with multiple locations (1-3).

When calculating the residuals we capped the residuals at the $90th$ percentile of the observed values (and set a minimum at the negative of this value) to reduce the influence of extreme observations. We calculated the standard error by a series of two delta transformations, converting the standard error of the log CDR first to linear space and secondly to logit space. Due to extreme values in these standard errors, we set bounds at the 25th and 75th percentiles of the standard errors.

For both steps we selected knots by taking evenly spaced quantiles of the dependent variable at the global level. For step 1 the knots fell at cumulative mobility disruptions of 0.16, 0.23, 0.30, and 0.40, and for step 2 the knots fell at 4, 6, and 8 (ie, April, June, and August).

We manually selected several other parameters to inform the spline models. Theta is a parameter that informs the strength of the prior when moving from one stage to the next in each step of the cascade. For step 1 we selected thetas of 4, 1, and 100 when moving between stages 1 and 2, stages 2 and 3, and stages 3 and 4, respectively. For step 2, we selected thetas of 20, 1, and 100. A smaller theta leads to a larger emphasis on the β from the earlier stage of the model in the estimation of β for the next stage. We also placed a prior on the maximum derivative of each spline segment with a Gaussian distribution with a mean of zero and standard error of one for step 1, stages 1-4. In step 2 we trimmed 10% of the data in the stage 1 model. In stages 2-4, we set a prior on the maximum derivative of each spline segment with a Gaussian distribution with a mean of zero and a standard error of 0.2. For the slope of the right linear tail in step 2 stages 2-4, we set a prior based on the slope value estimated in the step 2 stage 1 model: a

Gaussian distribution with a mean of -0.013 and a standard error of 0.2. This prior prevented extreme trends in the residuals in locations without data to support such trends.

We ran sensitivity analyses to examine the choice of thetas and strength of the Gaussian derivative prior. These results are provided in section 4.5.

In 94 countries with available data, we fit each step using the β estimates from each stage as a prior for the subsequent step, such that the estimates for the super-region drew strength from the global model, and the country-specific estimates drew strength from the regional and global models. The country-specific splines, therefore, only varied from the shape and magnitude of the global model if there was sufficient evidence to suggest a deviation.

In order to account for catch-up vaccination, we estimated all models in cumulative space from March 1, 2020. We chose March 1 because nearly every country experienced effects of COVID-19 on mobility by March. By using a *log* transformation and a monotonicity constraint in the step 1 model and a *logit* transformation in the step 2 model, we constrained the cumulative disruption between 0 and 1; therefore, in any one month, a country can exceed the expected rate of vaccine delivery (e.g. successfully implementing catch-up vaccination), but over time the estimated total vaccine delivery cannot exceed that expected in the absence of COVID-19.

Section 4.2: Uncertainty estimation

To calculate uncertainty intervals of disruption rates, we generated 1000 step 1 splines for each location based on the asymptotic posterior uncertainty intervals described in section 3.5 and a lognormal distribution. For countries with data, we generated a posterior set of 1000 splines based on asymptotic sampling from the mean and covariance matrix of the stage 3 splines. For countries without data, we randomly sampled with replacement from all 94 country-specific models to generate 1000 iterations. We included a sampling weight such that countries within the same super-region were 10 times as likely to be sampled as countries outside of the super-region.

For step 2, for 93 locations with nationally representative data, we similarly generated 1000 splines based on the stage 4 model, relying on asymptotic posterior uncertainty intervals. For countries without data, we generated 1000 splines from the super-region and vaccine-specific step 2 stage 3 model. Due to data sparsity $(N = 4)$ in central Europe, eastern Europe, and central Asia, for locations without data in this region, we relied on the step 2 stage 1 global model to estimate the residuals.

Based on the sum of these two sets of 1000 draws in logit space, we generated estimates based on the mean and provided 95% uncertainty intervals.

Section 4.3: In-sample validation

For each country, vaccine, and data source combination we calculated the average vaccine disruption ratio over the full time-period available to compare to the model results for the same location, vaccine, and time-period, by step. We calculated the weighted mean error (wME), weighted mean absolute error (wMAE), and root-mean-square error (RMSE) for each vaccine and step to evaluate the bias and goodness of fits. Step 1 involved modelling the relationship between cumulative disruptions in mobility and vaccine coverage, and step 2 incorporated residual variation. The size of each point is proportional to the inverse variance of the log ratio, the same weights used in fitting the model and calculating the wME and wMAE. The results are displayed in the figure and table below:

Due to the design of our model, predictions are constrained between zero and one. This aligns with our prior understanding that we do not believe any country would be likely to exceed expected coverage in the absence of COVID-19 under the conditions and challenges that the COVID-19 pandemic presented in 2020. We performed the same in-sample validation but held out all cumulative ratios over 1 and found that much of the bias (negative sign in the weighted mean error) is accounted for by CDRs over 1. The results are enumerated in the table below, only including CDRs < 1:

Section 4.4: Out-of-sample validation

We conducted two out-of-sample validation analyses to understand A) model performance when projecting to future months in locations with a partial time-series and B) model performance in locations without data. We used a stratified random sampling procedure to generate 20% holdouts, stratified by super-region.

In the first analysis, we aimed to estimate model performance when projecting in locations where we do not have a full time series (March-December) of data. First, as described above, we created a cross validation environment by generating 5 randomly selected holdout groups of countries, stratified by super-region. In each holdout group, we first withheld the last month of data from each country and ran a model. We then progressively held additional months of data, re-running models to produce out-ofsample estimates for each country. This produced 9 sets of out-of-sample validation models, withholding the terminal 1, 2, 3, …, and up to 9 months of data, respectively. For each country, we then compared the CDR from the terminal month in the full dataset to the out-of-sample predicted CDR for the same month after withholding *n* months. In this comparison, we only included countries for which we had at least 1 month of national-level data included in the model after holdouts. For example, 90 countries had nationallevel data for at least two months (March and April), so when dropping *n*=1 month of data, 90 countries were included in the comparison. Meanwhile, 43 countries had national-level data for at least 8 months (through at least October), so when dropping *n=*7 months of data, 43 countries were included in the comparison.

In the following plots, we show the number of countries (N), weighted mean error (wME), weighted mean absolute error (wMAE), and the root mean square error (RMSE) for each vaccine across the number of months withheld. We used the inverse variance of the log ratio as weights to calculate wMAE and wME. As expected, the wMAE are lowest when only withholding one month of data $(n = 1)$, but the results remain fairly consistent as *n* increases, even when we withhold the terminal 7+ months of data. This analysis suggests that our model performs well when projecting forward into time for countries where we only have a partial time-series of data included in the analysis.

As previously discussed in section 4.3, the design of the model constrains CDRs between zero and one. As in section 4.3, we performed a sensitivity analysis to evaluate the influence of this constraint on model fit, calculating out-of-sample wMAE, wME, and RMSE only for countries with CDRs < 1. This sensitivity analysis resulted in wME values closer to zero, suggesting that the negative bias is at least in part attributable to the presence of CDRs < 1 in the original dataset.

In the second analysis, we relied on the same stratified random sample of 20% holdouts but held out the entire time-series for each country in the group. This analysis serves to evaluate the performance of the model in locations where we have no data. Predictably, the model performance decreases in the absence of any data. This model is depicted in the figure below:

As the plot above illustrates, the out-of-sample model predictions tend to fall in a narrower range than the withheld data. Given the hierarchical nature of our model, this is to be expected: when no vaccination disruption data is available for a given country, the model uses the super-regional relationships between mobility and coverage disruption (from step 1 and step 2 of the model), along with country-specific mobility data, to predict disruptions. Within each out-of-sample model, all withheld countries for each super-region will therefore use these same super-regional relationships, with additional variation coming only from the withheld countries' mobility patterns. This process borrows strength within super-regions to produce predictions in countries without data, but as shown here may underrepresent true country-level variability in disruption. When aggregating to global or super-regional estimates, however, countries with under-estimated disruptions may be offset by countries with over-estimated disruptions during the aggregation process, as long as the model is generally unbiased.

We again repeated the same full-country-exclusion out-of-sample validation analysis excluding all CDRs over 1, as described for the partial-time-series out-of-sample validation analysis above and the in-sample validation analysis in section 4.3. Again, we find that much of the negative bias in this out-of-sample validation exercise can be attributed to the presence of CDRs > 1 in the original dataset. The results of this analysis are provided in the table below:

The majority of children eligible to receive doses of DTP3 and MCV1 live in countries with at least a partial time series of data (the scenario evaluated in the first of these out-of-sample validation analyses). Globally, two out of every three children in the target populations for each DTP3 and MCV1 live in one of the 94 countries that have data included in this analysis. The following table details the fraction of the target population for each antigen who live in countries with data included in the analysis, globally and by GBD super-region:

Section 4.5: Spline parameter sensitivity analysis

We ran the following sensitivity analyses to understand the impact of the spline parameters on the results:

As shown in this table, the global results were only somewhat sensitive to these model parameters. Though some settings (weaker influence of global and regional models and more flexible splines) improved in-sample validation criteria (wME and wMAE), increased model flexibility raises the risk of overfitting available data. The values selected in the manuscript analysis balanced our prior expectations and global trends with location-specific results to ensure outliers did not overly influence the model.

The final row of the table presents a sensitivity analysis conducted to examine the potential effects of bias due to reporting delays in the last month of administrative data reported to the WHO regional offices. In the primary analysis, we attempt to limit the effects of this bias by excluding particular countries with evidence of large decreases (greater than 20%) in the last month of available data (section 2.1). To understand the impact of this strategy and to understand the potential impact of reporting delays, we reran the analysis dropping all data from the last month reported to the WHO regional offices. The results were fairly robust to this change with an increase in the global annual disruption of about 1% for both antigens; uncertainty intervals for the primary (manuscript) analysis and sensitivity analysis broadly overlapped. Eliminating the last month of data for all countries in this sensitivity analysis systematically excludes data points that suggest ongoing catch-up vaccination and/or recovery of services in the latter months of 2020. This sensitivity analysis therefore results in a larger disruption estimate (and potential over-estimation of the true magnitude of disruption) compared to the primary analysis presented in the manuscript.

Section 4.6: GBD vaccine coverage estimates $¹$ </sup>

Annual estimates of childhood routine vaccination coverage, including DTP3 and MCV1 are produced as part of the larger Global Burden of Diseases, Injuries, and Risk Factors Study (GBD).²³⁻²⁵ The methodology for producing these estimates has been described elsewhere in detail and is adapted below.¹

Vaccination coverage was defined as "the proportion of children that received at least the stated number of doses (e.g. DTP1) from a routine immunisation delivery system."1 The GBD estimates leveraged survey microdata from children ages 12 months through 59 months, survey report data in the absence of microdata, and bias-adjusted official country-reported estimates available from the 2020 Joint Reporting Form (JRF).¹⁵ We excluded doses delivered through vaccination campaigns whenever possible. Data was assigned to year of expected delivery using national immunisation schedules from the JRF,²⁶ assuming negligible effects of any catch-up vaccination, migration, or differential mortality in older age cohorts. In addition to direct coverage data, data on country-reported stockouts or disruptions to routine delivery were also incorporated.¹⁵ These data were used to compute a covariate used in modelling which captured the expected magnitude of delivery disruption proportional to the change in bias-adjusted official countryreported data in years when a reported stockout was matched by a decrease in reported coverage.

Spatio-temporal Gaussian process regression (ST-GPR) was the primary modelling tool used for this analysis. Details on ST-GPR estimation for the GBD has been described elsewhere.24 In brief, ST-GPR is a stochastic modelling framework designed to incorporate diverse input data and predict over space and time. It uses a selected linear regression as the basis of every model, followed by locally-weighted regression (LOESS) to smooth and minimise prediction error, then Gaussian process regression (GPR) to predict smoothed estimates over space and time while accounting for underlying data uncertainty. DTP3 and MCV1 were informed by the following linear model:

$$
logit(P_c(t)) = \beta_0 + \beta_1 HAQ_{c,t} + \beta_2 war_{c,t} + \beta_3 stockout_{c,t} + \alpha_c + \gamma_{R[c]} + \omega_{SR[c]} + \epsilon_{c,t}
$$

where $P_c(t)$ is vaccination coverage for country c year t; $HAQ_{c,t}$ is value of the health care access and quality (HAQ) index²⁷ for country c and year t; war_{c,t} is log-transformed mortality rate due to acute war and terror events; stockout_{c,t} is value of the vaccine stockout covariate for country c and year t; and α_c , $\gamma_{R[c]}$, and $\omega_{SR[c]}$ are nested country, region, and super-region random intercepts, respectively.

From the DTP3 and MCV1 models, respectively, we obtained 1,000 samples (draws) from the posterior distribution of coverage. Draws were summarised to produce mean estimates of annual vaccination coverage complemented by 95% uncertainty intervals based on the $2.5th$ and 97.5th ordinal draws. Estimates were produced globally for 204 countries and territories, 1980 to 2019.

Section 4.7: Estimation of expected 2020 coverage in the absence of COVID-19

Using the ST-GPR estimation framework described in section 4.6 above, we assumed that in the absence of COVID-19, country-specific national coverage trends estimated from 1980-2019 were expected to continue in the absence of the pandemic. In order to produce these estimates, we fit ST-GPR models to available vaccine coverage data from 1980-2019 as described above. We then used the fitted models to produce predictions for 2020, using modeled estimates of the HAQ index in 2020 – representing expected values of HAQ in the absence of the COVID pandemic – as a covariate. This process produced a set of estimates of vaccine coverage in the absence of COVID-19 for 2020, assuming that past trends in coverage (and HAQI) would have continued into the future. In the 2020 counterfactual, we did not include effects of any national-level vaccine stockouts nor disruptions due to war and terror given the absence of available data.

References

- 1 Galles NC, GBD 2020 Vaccine Coverage Collaborators. Measuring routine childhood vaccination coverage in 204 countries and territories, 1980-2019: a systematic analysis for the Global Burden of Disease Study 2020, Release 1. *The Lancet* 2021; published online Feb.
- 2 Zheng P, Barber R, Sorensen R, Murray C, Aravkin A. Trimmed Constrained Mixed Effects Models: Formulations and Algorithms. *J Comput Graph Stud* 2021; **0**: 1–13.
- 3 Moreno Pérez D. Vacunas en todas las edades: Que el COVID-19 no nos frene. 2020; published online June 2. https://www.vacunas.org/webinar-vacunas-en-todas-las-edades-que-el-covid-19-no-nos-frene/ (accessed Feb 5, 2021).
- 4 Department of Health Services, Ministry of Health and Population. Nepal Monthly Immunization Coverage by Province 2019-2020. Nepal: Department of Health Services, Ministry of Health and Population.
- 5 Ministry of Health and Family Welfare. India HMIS Standard Report Data Itemwise Monthly 2018- 2021. New Delhi, India: Ministry of Health and Family Welfare https://nrhmmis.nic.in/hmisreports/frmstandard_reports.aspx.
- 6 National Centre for Immunization Research and Surveillance. COVID-19: Impact on routine childhood vaccination uptake in Australia. Sydney, Australia: National Centre for Immunization Research and Surveillance, 2020 https://ncirs.org.au/sites/default/files/2020-11/COVID-19_Impact_Analysis_Final%20Report.pdf (accessed Nov 30, 2020).
- 7 National Health Analytical Tool showing Health Service Uptake during the COVID-19 pandemic. Nigerian National Health Management Information System https://public.tableau.com/views/COVID19AnalysisofHealthServiceUptakeusingNHMISData/MAIND ashboard (accessed March 1, 2021).
- 8 Public Health England. Impact of COVID-19 on childhood vaccination counts up to week 11, 2021. London, England: Public Health England, 2021 https://www.gov.uk/government/publications/covid-19-impact-on-vaccination-programmes/impact-of-covid-19-on-childhood-vaccination-counts-up-toweek-11-2021.
- 9 COVID-19 Wider Impacts Tool. Public Health Scotland https://scotland.shinyapps.io/phs-covidwider-impact/ (accessed March 1, 2021).
- 10 Daley J. Vaccinations Have Sharply Declined Nationwide during the COVID-19 Pandemic. Sci. Am. https://www.scientificamerican.com/article/vaccinations-have-sharply-declined-nationwide-duringthe-covid-19-pandemic/ (accessed Aug 13, 2020).
- 11 Masters K, April 27 VM, 2020. Virginia pediatricians say reported decline in vaccination rates amid COVID-19 pandemic is troubling. Va. Mercury. 2020; published online April 27. https://www.virginiamercury.com/2020/04/27/virginia-pediatricians-report-troubling-decline-invaccination-rates-as-covid-19-pandemic-continues/ (accessed Nov 30, 2020).
- 12 UNICEF, World Health Organization. Progress and Challenges with Achieving Universal Immunization Coverage: 2019 WHO/UNICEF Estimates of National Immunization Coverage [Unpublished Data]. UNICEF/WHO, 2020 https://www.who.int/immunization/monitoring_surveillance/who-immuniz.pdf.
- 13 Lassi ZS, Naseem R, Salam RA, Siddiqui F, Das JK. The Impact of the COVID-19 Pandemic on Immunization Campaigns and Programs: A Systematic Review. *Int J Environ Res Public Health* 2021; **18**. DOI:10.3390/ijerph18030988.
- 14 Danovaro MC. The impact of COVID-19 on immunization programmes publications. TechNet-21. https://www.technet-21.org/en/topics/covid-19-impact (accessed May 11, 2021).
- 15 WHO/UNICEF Joint Reporting Process. WHO http://www.who.int/immunization/monitoring_surveillance/routine/reporting/en/ (accessed April 8, 2021).
- 16 World Health Organization. Special feature: Immunization and COVID-19: Second pulse poll to help understand disruptions to vaccination and how to respond. Geneva, Switzerland: WHO Global Immunization News (GIN) Newsletter, 2020.
- 17 World Health Organization. Pulse survey on continuity of essential health services during the COVID-19 pandemic. Geneva, Switzerland: World Health Organization, 2020.
- 18 Danovaro C, Gurung S, Ho LL, Linstrand A. Understanding the disruption to programmes through rapid polling. WHO Global Immunization News (GIN) Newsletter, 2020.
- 19 Lee CH, Cook S, Lee JS, Han B. Comparison of Two Meta-Analysis Methods: Inverse-Variance-Weighted Average and Weighted Sum of Z-Scores. *Genomics Inform* 2016; **14**: 173–80.
- 20 Bland JM, Altman DG. The odds ratio. *BMJ* 2000; **320**: 1468.
- 21 Friedman JH. Multivariate Adaptive Regression Splines. *Ann Stat* 1991; **19**: 1–67.
- 22 Boor C de, DeBoor C. A Practical Guide to Splines. Springer New York, 1978.
- 23 Vos T, Lim SS, Abbafati C, *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 2020; **396**: 1204–22.
- 24 Murray CJL, Aravkin AY, Zheng P, *et al.* Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 2020; **396**: 1223–49.
- 25 Wang H, Abbas KM, Abbasifard M, *et al.* Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *The Lancet* 2020; **396**: 1160–203.
- 26 World Health Organization. WHO and UNICEF Reported Immunization Schedules 2000-2019. Geneva, Switzerland: World Health Organization, 2020.
- 27 Fullman N, Yearwood J, Abay SM, *et al.* Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *The Lancet* 2018; **391**: 2236–71.

Supplementary results

Comparing vaccine and mobility patterns

The global step 1 model of cumulative vaccine disruption versus cumulative mobility disruption suggested steep declines in vaccine coverage at initial mobility decreases and continued declines through the entire range of cumulative mobility disruption. Data suggested reductions in vaccine coverage even at low levels of mobility reduction. In step 1, we found a wide range of vaccine disruption ratios across countries with similar mobility disruptions. At a 40% cumulative disruption in mobility, DTP3 vaccine disruption estimates across locations ranged from 20.0% of expected to no disruption, with a standard deviation of 16.6%. For MCV1, disruption estimates ranged from 17.7% of expected to no disruption, with a standard deviation of 18.1%.

The global step 2 model of residuals over time, suggest that recovery in vaccine delivery is outpacing that which would be expected from the relationship with mobility alone. Though the model does not distinguish doses delivered to children missed early in the pandemic from those delivered on schedule as services resume, in later months, positive residuals show that disruptions tend to be smaller compared the disruptions predicted by mobility trends. This suggests that active efforts to mitigate RI disruptions may be helping to accelerate the pace of recovery in many locations.

Supplementary tables

Supplementary table 1. Data sources for vaccine delivery disruption and metadata. AL=Alabama. AR=Arkansas. CA=California. CDR=Cumulative Disruption Ratio. CO=Colorado. DHIS-2=District Health Information Software 2. DTP=diphtheria, tetanus, and pertussis. DTP3=diphtheria, tetanus, and pertussis, third dose. FL=Florida. HMIS=Health Management Information System. IA=Iowa. ID=Idaho. IL=Illinois. KS=Kansas. KY=Kentucky. MA=Massachusetts. MCV1=measles-containing vaccine, first dose. MI=Michigan. MMR=measles, mumps, and rubella. MMR1=measles, mumps, and rubella, first dose. MN=Minnesota. MR=measles and rubella. NCIRS=National Centre for Immunisation Research and Surveillance. ND=North Dakota. NE=Nebraska. NHMIS=National Health Management Information System. NJ=New Jersey. NV=Nevada. OK=Oklahoma. OR=Oregon. PCV3=pneumococcal conjugate vaccine, third dose. Penta3=pentavalent vaccine, third dose; covers diphtheria, tetanus, pertussis, hepatitis B, and *Haemophilus influenzae* type b. RI=Rhode Island. SC=South Carolina. SD=South Dakota. TN=Tennessee. TX=Texas. USA=United States of America. UT=Utah. VA=Virginia. VT=Vermont. WI=Wisconsin.

Supplementary table 2. Expected DTP3 and MCV1 coverage in the absence of COVID-19 and estimates of disruptions and coverage attributable to the pandemic in 2020, by country. Expected coverage was expected levels for 2020 in the absence of the COVID-19 pandemic based on past trends. Estimated coverage reflects coverage for 2020 while accounting for estimated pandemic-related disruptions; more detail on these methods is in the main manuscript and appendix section 4. DTP3=diphtheria, tetanus, and pertussis, third dose. MCV1=measles-containing vaccine, first dose. UI=uncertainty interval.

AL=Alabama. AR=Arkansas. CA=California. CDR=Cumulative Disruption Ratio. CO=Colorado. DHIS-2=District Health Information Software 2. DTP=diphtheria, tetanus, and pertussis. DTP3=diphtheria, tetanus, and pertussis, third dose. FL=Florida. HMIS=Health Management Information System. IA=Iowa. ID=Idaho. IL=Illinois. KS=Kansas. KY=Kentucky. MA=Massachusetts. MCV1=measles-containing vaccine, first dose. MI=Michigan. MMR=measles, mumps, and rubella. MMR1=measles, mumps, and rubella, first dose. MN=Minnesota. MR=measles and rubella. NCIRS=National Centre for Immunisation Research and Surveillance. ND=North Dakota. NE=Nebraska. NHMIS=National Health Management Information System. NJ=New Jersey. NV=Nevada. OK=Oklahoma. OR=Oregon. PCV3=pneumococcal conjugate vaccine, third dose. Penta3=pentavalent vaccine, third dose; covers diphtheria, tetanus, pertussis, hepatitis B, and Haemophilus influenzae type b. RI=Rhode Island. SC=South Carolina. SD=South Dakota. TN=Tennessee.
TX=Texas. USA=United S

Expected coverage was expected levels for 2020 in the absence of the COVID-19 pandemic based on past trends. Estimated coverage reflects coverage for 2020 while accounting for estimated pandemic-related disruptions; more detail on these methods is in the main manuscript and appendix section 4. DTP3=diphtheria, tetanus, and pertussis, third dose. MCV1=measles-containing vaccine, first dose. UI=uncertainty interval.