Statistical Analysis

Percy Soto-Becerra

## Statistical Analysis (main article)

This study employed an interrupted time-series (ITS) analysis approach, using segmented regression, to appraise the impact of the COVID-19 pandemic on mental health service utilisation rates. The data for each geographical context were univariate and monthly in nature, necessitating the fitting of separate models for each country. Supplementary table SX provides an overview of the start and end dates of each time series for each country and mental health category, the interruption date considered for each country, and any pre-pandemic changes to the registration system. Notably, for China, Norway, and the USA, post-pandemic changes in the registration system led us to censor the analysis up until the month before these shifts. For China, a pre-pandemic change was explicitly modelled using a dummy variable.

Our primary analysis incorporated a washout period for March 2020, considering the onset of the pandemic to be April 2020. We also conducted sensitivity analyses to account for potential delayed effects and the absence of a washout period.The ITS analysis comprised two segments: the counterfactual model, representing the expected trend in the absence of the pandemic, and the impact model, examining the immediate and sustained effects of the pandemic. Interpretation of the level change as an immediate effect will reflect the instantaneous change in mental health service utilisation following the pandemic’s onset. Concurrently, the trend change will represent the sustained effect, illustrating any alterations in the outcome variable’s slope post-intervention compared to pre-intervention. The segmented regression strategy facilitated the derivation of relative rates along with their respective 95% confidence intervals and p-values, obtained using the Wald method.

To handle potential over-dispersion in the count data, we employed a generalised linear mixed model (GLMM) with penalised quasi-likelihood (PQL) estimation. An autoregressive moving average (ARMA) structure was fitted to the normalised Pearson residuals to manage autocorrelation, a crucial consideration for maintaining the robustness of our statistical inferences.

Several diagnostic checks were performed to ensure the validity of the analysis. The linearity and potential heteroscedasticity were assessed by examining plots of fitted values versus normalised Pearson residuals. Normalised Pearson residuals, as opposed to conventional Pearson residuals, were used in order to account for the autocorrelation already corrected for in the model, thereby providing a more rigorous and robust model validation. Normal quantile-quantile plots of these residuals were scrutinised to evaluate the normality assumption of the residuals. Even though the response variable follows a Poisson distribution, this normality check of residuals is important as it verifies the adequacy of the PQL approximation used in the model fitting. The assumption of independent errors was confirmed by analysing autocorrelation plots of these residuals. Lastly, the model’s predictive accuracy was verified by investigating calibration plots of predicted versus observed counts.

Supplementary table SX provides an overview of the start and end dates of each time series for each country and mental health category, the interruption date considered for each country, and any pre-pandemic changes to the registration system. Notably, for China, Norway, and the USA, post-pandemic changes in the registration system led us to censor the analysis up until the month before these shifts. For China, a pre-pandemic change was explicitly modelled using a dummy variable.

Further details regarding the sensitivity analyses are summarised in supplementary table SX2. All analyses were conducted using R version 4.3.0.

## Supplementary Methods

### Detailed model’s description

The outcome variable is the incidence rate for diagnosis in month . Our model is:

Here: - is the log of the incidence rate at the onset of the pandemic. - is the linear time effect before the pandemic. - is an indicator variable for the pandemic (1 if after April 2020, 0 otherwise). - is the immediate change in incidence rate at the onset of the pandemic. - is the change in slope (sustained effect on the incidence rate) post-pandemic. - are sine and cosine transformations to control seasonality. - is an indicator variable for changes in the registry system in China. - and are random effects for country , with , where is the variance-covariance matrix of the random effects. - are the error terms, which are assumed to follow an ARMA(, ) model for residual autocorrelation.

The harmonic transformations are:

The autocorrelation is described by:

Here, is the white noise error term, are the autocorrelation coefficients, and are the moving average coefficients. These coefficients are selected based on autocorrelation and partial autocorrelation plots.