

Background

This methodological note details the steps adopted for the progress analysis of the sustainable development indicators, corresponding to Goal 3 of the SDGs. The Sustainable Development Goals group 17 goals comprising 169 targets. Goal 3 focuses on *ensuring healthy lives and promoting well-being for all at all ages*. Many health indicators were moving in the right direction prior to the emergence of COVID-19.

On progress toward meeting targets, maternal and child health has improved, immunization coverage has increased, and communicable diseases have declined, although not fast enough to meet the 2030 compact goals. The pandemic has caused efforts to keep the population healthy to focus on the emerging situation and has slowed and even reversed progress in health, posing a major threat beyond the disease itself. Some 90% of countries continue to report one or more interruptions in essential health services and available data from some countries show that the pandemic has shortened the life expectancy of populations. Not surprisingly, this systemic effect disproportionately affects disadvantaged groups.

The pandemic has demonstrated the importance of universal health coverage and multisectoral coordination for health emergency preparedness. In addition, to design effective pandemic policies, governments will need to improve and strengthen the collection of basic demographic and epidemiological data. Therefore, an assessment of progress under SDG3 is essential to measure the challenges that still need to be overcome. Thus, to measure progress in terms of meeting the agreed targets for SDG3, analyses were conducted to compare the progress shown by countries against what would be expected in terms of meeting the targets agreed at the United Nations. Therefore, 5 analytical steps were taken to characterize the situation of the Americas with respect to the achievement of the goals related to SDG3:

1. Analysis of the evolution of the average of the indicators at the regional level over time;
2. The estimated time required to reach the 2030 target;
3. The temporal evolution of social inequality between countries in the indicator;
4. The forecast of the indicator's temporal progress; and,
5. The magnitude of the changes observed in the averages and inequalities of the indicator.

All the structuring of the analyses was documented to ensure reproducibility. Thus, all extraction, transformation and loading (ETL) routines were performed using structured code with the R statistical programming language. Microsoft Power BI was used to generate visualizations of the metrics linked to the analysis results.

Data sources

The data used to characterize the situation of the countries and corresponding subregions were extracted from the Global Health Observatory (GHO)¹ of the World Health Organization (WHO). The observatory's API provides a simple interface for querying data content and statistics from the World Health Organization, using OData (Open Data Protocol). Following the data request standards, an R routine has been structured to extract the latest available version of the 2,195 indicators available in the GHO OData API. The code used for this process can be found here: [GHO OData API](#). From the list of 2,195 indicators, data corresponding to the SDG3 indicators for countries or territories in the Americas region were selected. The following table highlights the status of each indicator in terms of data availability, as well as the codes used.

Table 1 - Data availability with the GHO OData API.

| Indicator | State | API Indicator Code | An alternative version is being studied |
|---|-----------|----------------------------------|---|
| SDG 3.1.1 - Maternal Mortality Ratio | Found | MDG_0000000026 | N/A |
| SDG 3.1.2 - Births attended by skilled health personnel | Found | MDG_0000000025 | N/A |
| SDG 3.2.1 - Under-5 mortality rate | Found | u5mr | N/A |
| SDG 3.2.2 - Neonatal Mortality Rate | Found | nmr | N/A |
| SDG 3.3.1 - Number of new HIV infections | Found | HIV_0000000026 | N/A |
| SDG 3.3.2 - Tuberculosis Incidence | Found | MDG_0000000020 | N/A |
| SDG 3.3.3 - Estimated incidence of malaria | Found | MALARIA_EST_INCIDENCE | N/A |
| SDG 3.3.4 - Incidence of hepatitis B per 100,000 inhabitants | Not found | WHS4_117 | Hepatitis B surface antigen prevalence - children under 5 years of age. |
| SDG 3.3.5 - Number of interventions against NTDs | Found | SDGNTDTREATMENT | N/A |
| SDG 3.4.1 - Age-standardized NCD mortality rate | Found | NCDMORT3070 | N/A |
| SDG 3.4.2 - Age-standardized suicide rates | Found | MH_12 | N/A |
| SDG 3.5.1 - Coverage of treatment interventions (pharmacological, psychosocial, rehabilitation and aftercare services) for substance use disorders. | Not found | N/A | N/A |
| SDG 3.5.2 - Alcohol, harmful use (15+), 12-month prevalence | Found | SA_0000001754 | N/A |
| SDG 3.6.1 - Estimated road traffic fatality rate | Found | RS_198 | N/A |
| SDG 3.7.1 - Women with modern family planning methods | Found | SDGFPALL | N/A |
| SDG 3.7.2 - Adolescent birth rate per 1,000 women | Found | MDG_0000000003 | N/A |
| SDG 3.8.1 - Coverage of essential health services | Found | UHC_INDEX_REPORTED | N/A |
| SDG 3.8.2 - Proportion of population with high household expenditure on health in relation to total household expenditure or income. | Not found | N/A | N/A |
| SDG 3.9.1 - Mortality rate from ambient and domestic air pollution | Found | SDGAIRBOD | N/A |
| SDG 3.9.2 - Death rate attributed to unsafe WASH | Found | SDGWASHBOD | N/A |
| SDG 3.9.3 - Mortality rate from unintentional poisoning | Found | SDGPOISON | N/A |
| SDG 3.a.1 - Estimating current tobacco use prevalence | Found | M_Est_tob_curr_std | N/A |
| SDG 3.b.1 - Proportion of the target population covered by all vaccines included in their national program. | Not found | N/A | N/A |
| SDG 3.b.2 - Official net development of basic health/medical research | Found | SDGODA | N/A |
| SDG 3.b.3 - Health facilities with a basic set of essential drugs | Found | SDGHEALTHFACILITIESESSENTIALMEDS | N/A |
| SDG 3.c.1 - Density of qualified health professionals | Found | SDGSHP | N/A |
| SDG 3.d.1 - International Health Regulations (IHR) capacity and preparedness for health emergencies | Not found | N/A | Basic capacity of the International Health Regulations |

¹ The documentation for the API query terms can be found here: <https://www.who.int/data/gho/info/gho-odata-api>.

SDG 3.d.2 - Percentage of bloodstream infections due to selected antimicrobial resistant organisms.

Not found

N/A

N/A

The indicators listed below were not found in the GHO OData API as described in the SDG3 targets, so where available, an alternative version was considered.

- SDG-3.3.4 - Incidence of hepatitis B per 100,000 population - Alternative considered: prevalence of hepatitis B surface antigen (HBsAg) among children under 5 years of age.
- SDG -3.5.1 Coverage of treatment interventions (pharmacological, psychosocial, rehabilitative and aftercare services) for substance use disorders - Alternative considered: None
- SDG -3.8.1 Essential health services coverage (defined as the average coverage of essential services based on screening interventions including reproductive, maternal, newborn and child health, infectious diseases, non-communicable diseases, and capacity and access to services, among the general population and the most disadvantaged) - Alternative considered: CSU Services Coverage Index.
- SDG -3.8.2 Proportion of population with significant household health expenditures as a proportion of total household income or expenditure - Alternative considered: None.
- SDG -3.b. 1 Proportion of the target population covered by all vaccines included in their national program - Alternative considered: None
- SDG -3.d. 2 Percentage of bloodstream infections due to selected antimicrobial resistant organisms - Alternative considered: None

In order to make a decision on the data source to be used for the progress assessment, a missing values analysis was performed in relation to the data provided by GHO. Taking into account the situation regarding the availability of the ODS3 data, we conducted a quality analysis of the missing information in relation to the indicators that were available in GHO's OData API. We found that the existence of the indicator in the API repository is not associated with perfect availability of the information. Countries share data irregularly. In the following graphs we analyzed the quality of the information by indicator and by country. In this way, it was possible to select which indicators and countries can be considered for further analysis and automation. The indicator data spans from 2000 to the latest date available in the GHO OData API. Each value marked as "missing" represents missing information for an observation point. Thus, indicators with the column completely shaded indicate total missing information with the GHO OData API, such as SDG 3.5.2. The higher the number of shaded lines, the more missing values were observed.

Figure 1 - Analysis of missing values by indicator.

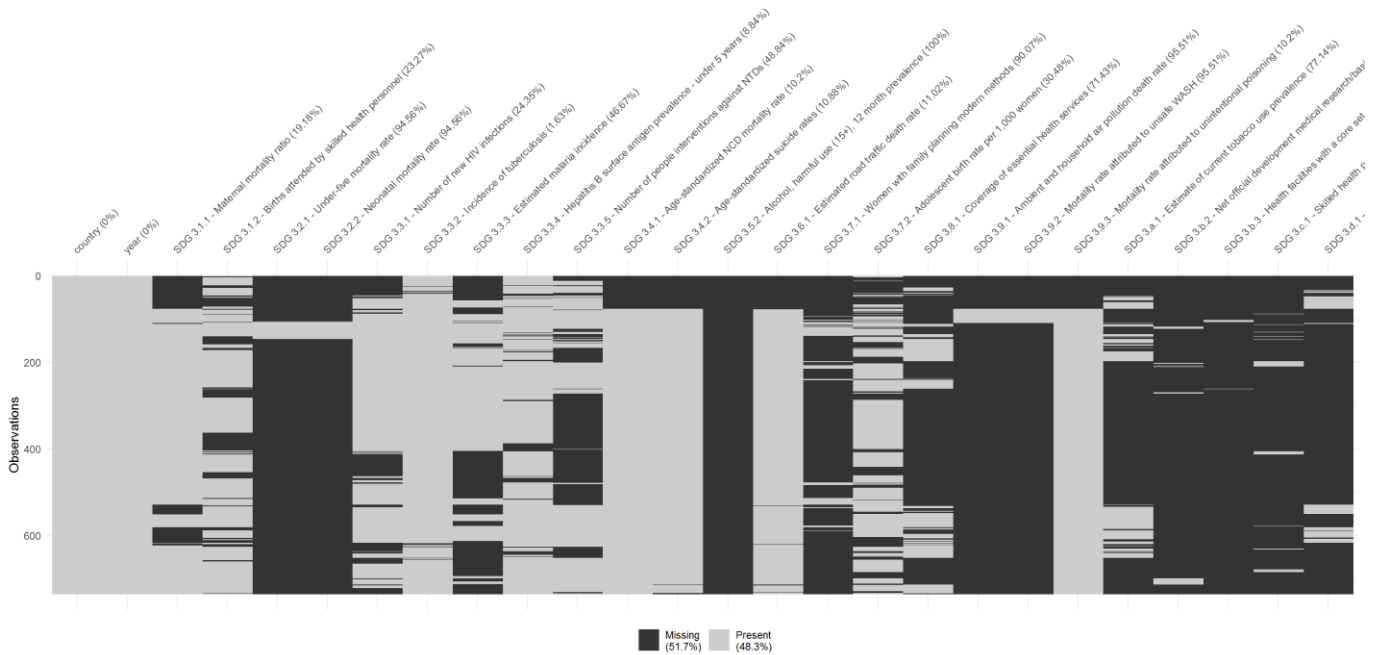
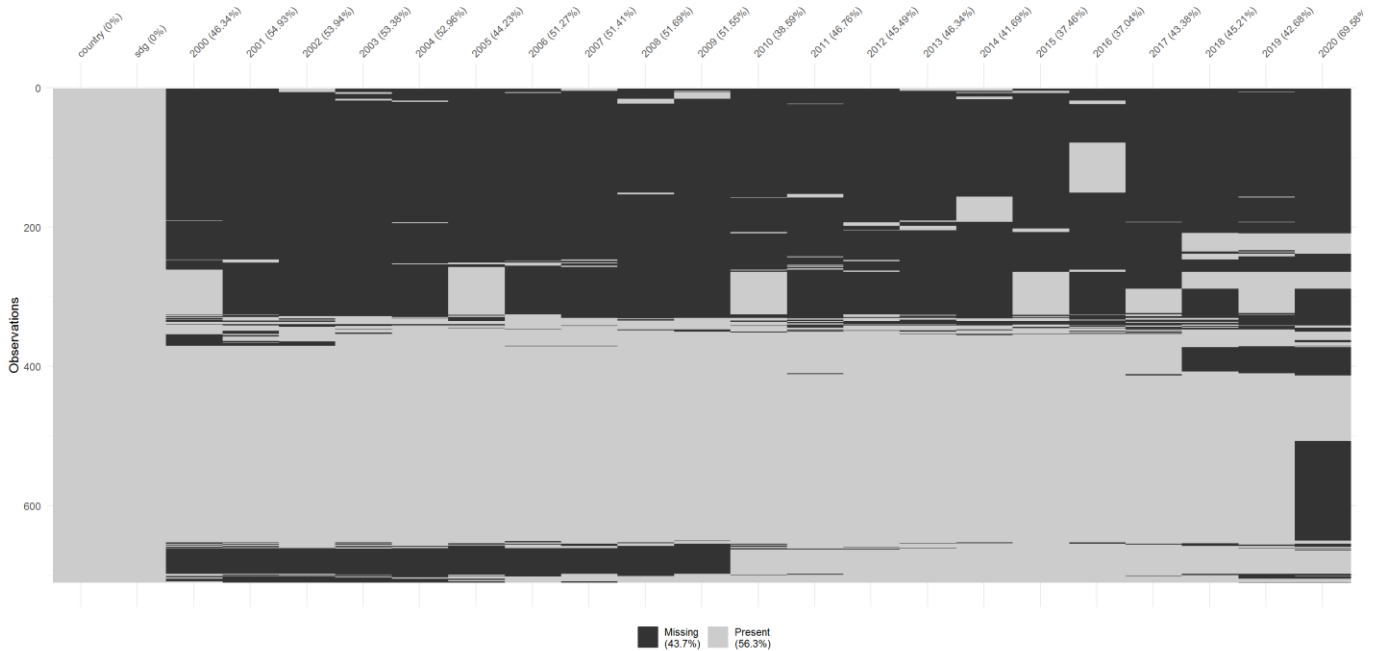


Figure 2 - Analysis of missing values by time.



For some indicators, they could not be found in the GHO API. To collect infant and child mortality data, for example, we chose to use IGME (United Nations Interagency Group for Child Mortality Estimation) as a source, due to the unavailability of these data in the GHO. In the Table 2 details

the respective data source for each of the indicators. It is worth mentioning that for several indicators there is no systematized data source, with records in time series format, that would allow for progress evaluations. These indicators have not been considered for analysis purposes in this report due to this limited availability of information.

Table 2 - SDG3 indicators and their respective data sources.

| ODS | Indicator detail | Source of data | Situation in relation to reports | Time scope |
|-----------|---|---|------------------------------------|--------------------------------|
| SDG 3.1.1 | Maternal mortality rate | Trends in maternal mortality 2000 to 2020: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/Population Division | Considered | 2000-2020 |
| SDG 3.1.2 | Deliveries attended by skilled health personnel | GHO | Considered | 2000-2019 |
| SDG 3.2.1 | Under-five mortality rate | IGME | Considered | 2000-2021 |
| SDG 3.2.2 | Neonatal mortality rate | IGME | Considered | 2000-2021 |
| SDG 3.3.1 | Number of new HIV infections | AIDSinfo | Considered | 2000-2020 |
| SDG 3.3.2 | Incidence of tuberculosis | Global Tuberculosis Report 2022 | Considered | 2000-2021 |
| SDG 3.3.3 | Estimated incidence of malaria | GHO | Considered | 2000-2019 |
| SDG 3.3.4 | Hepatitis B surface antigen prevalence - children under 5 years of age. | GHO | Considered | 2000-2019 |
| SDG 3.3.5 | Number of people intervened against the ETD | GHO | Considered | 2010-2019 |
| SDG 3.4.1 | Age-standardized NCD mortality rate | GHO | Considered | 2000-2019 |
| SDG 3.4.2 | Age-standardized suicide rates | GHO | Considered | 2000-2019 |
| SDG 3.5.1 | Coverage of substance use disorder treatment interventions. | GHO | Not considered - insufficient data | - |
| SDG 3.5.2 | Harmful use of alcohol (15+) 12-month prevalence | GHO | Not considered - insufficient data | - |
| SDG 3.6.1 | Estimated mortality rate due to traffic accidents | GHO | Considered | 2000-2019 |
| SDG 3.7.1 | Women with modern family planning methods | GHO | Not considered - insufficient data | - |
| SDG 3.7.2 | Adolescent birth rate per 1,000 females | GHO | Considered | 2000-2019 |
| SDG 3.8.1 | Coverage of essential health services | GHO | Evaluate | 2000,2005,2010, 2015,2017,2019 |
| SDG 3.8.2 | Proportion of large household expenses 10%. | GHO | Not considered - insufficient data | - |
| SDG 3.8.2 | Proportion of major household expenses 25% of total household expenses | GHO | Not considered - insufficient data | - |

| ODS | Indicator detail | Source of data | Situation in relation to reports | Time scope |
|------------|---|----------------|------------------------------------|--------------------------------|
| SDG 3.9.1 | Mortality rate due to environmental and domestic air pollution | GHO | Not considered - insufficient data | 2016 |
| SDG 3.9.2 | Mortality rate attributed to unsafe WASH | GHO | Not considered - insufficient data | 2016 |
| SDG 3.9.3 | Mortality rate attributable to unintentional poisoning | GHO | Considered | 2000-2019 |
| SDG 3.a. 1 | Estimation of current prevalence of tobacco use | GHO | Considered | 2000,2005,2010, 2015,2018,2019 |
| SDG 3.b. 2 | Official development net medical research basic health care | GHO | Not considered - insufficient data | - |
| SDG 3.b. 3 | Health centers with a basic set of essential medicines | GHO | Not considered - insufficient data | - |
| SDG 3.c. 1 | Density of qualified health professionals | GHO | Not considered - insufficient data | - |
| SDG 3.d. 1 | Basic capacity of the International Health Regulations | GHO | Not considered - insufficient data | - |
| SDG 3.d. 2 | Percentage of bloodstream infections due to antimicrobial resistant organisms | GHO | Not considered - insufficient data | - |

Establishment of objectives to evaluate progress

All the data obtained represented information from the countries and territories of the Americas region. However, the analyses were conducted by subregion. The aggregation of country data into subregions was done in order to show more consolidated trends in relation to progress in achieving the goals linked to SDG3, as the regional data would be less subject to fluctuations in the values presented. Therefore, a recalculation methodological procedure was adopted to estimate the indicators for each subregion using weighted averages. Each indicator was broken down to obtain the absolute number of events related to that indicator in each country. Once this number was obtained, all the values for the countries in each subregion were added together. This sum was weighted by the population denominator recommended for each indicator, taking into account the subregional sum. Thus, it was possible to recalculate the value of the indicators by aggregating the data for each country by subregion. All the analyses and graphs presented in this report followed this methodology, with the exception of the equity metrics, where country data were analyzed for each of the subregions considered without the regional recomposition detailed above. For the analyses by subregion, the categorization detailed in Table 3.

Table 3 - Grouping of countries by subregion.

| Country | Subregion |
|------------------------------------|---------------------|
| Bermuda shorts | North America |
| Canada | North America |
| United States of America | North America |
| Brazil | Brazil |
| Cuba | Latin Caribbean |
| Dominican Republic | Latin Caribbean |
| Guadalupe | Latin Caribbean |
| French Guiana | Latin Caribbean |
| Haiti | Latin Caribbean |
| Martinique | Latin Caribbean |
| Puerto Rico | Latin Caribbean |
| Aruba | Non-Latin Caribbean |
| Eel | Non-Latin Caribbean |
| Antigua and Barbuda | Non-Latin Caribbean |
| Bahamas | Non-Latin Caribbean |
| Barbados | Non-Latin Caribbean |
| Curaçao | Non-Latin Caribbean |
| Cayman Islands | Non-Latin Caribbean |
| Dominica | Non-Latin Caribbean |
| Grenada | Non-Latin Caribbean |
| Guyana | Non-Latin Caribbean |
| Jamaica | Non-Latin Caribbean |
| St. Kitts and Nevis | Non-Latin Caribbean |
| St. Lucia | Non-Latin Caribbean |
| Montserrat | Non-Latin Caribbean |
| Suriname | Non-Latin Caribbean |
| Sint Maarten (Netherlands) | Non-Latin Caribbean |
| Turks and Caicos Islands | Non-Latin Caribbean |
| Trinidad and Tobago | Non-Latin Caribbean |
| Saint Vincent and the Grenadines | Non-Latin Caribbean |
| British Virgin Islands | Non-Latin Caribbean |
| Virgin Islands (U.S.) | Non-Latin Caribbean |
| Belize | Central America |
| Costa Rica | Central America |
| Guatemala | Central America |
| Honduras | Central America |
| Nicaragua | Central America |
| Panama | Central America |
| El Salvador | Central America |
| Argentina | Southern Cone |
| Chile | Southern Cone |
| Paraguay | Southern Cone |
| Uruguay | Southern Cone |
| Mexico | Mexico |
| Bolivia (Plurinational State of) | Andean zone |
| Colombia | Andean zone |
| Ecuador | Andean zone |
| Peru | Andean zone |
| Venezuela (Bolivarian Republic of) | Andean zone |

The assessment of progress took into account those indicators associated with SDG3 that had targets, either stipulated by agreement in relation to the targets agreed between the United Nations and the member countries, or in relation to the 2030 Agenda for health in the Americas. The choice of using a combination of targets between the goals established by the United Nations and the 2030 Agenda was due to the existence of targets already achieved for the region of the Americas, when taking into account only the results defined by the United Nations with the SDG3. Thus, the 2030 Agenda targets, for some indicators, reflect a more objective measurement of progress in improving health in the region.

With regard to the monitoring of SDG 3 targets, four possibilities were considered for their definition:

1. Targets in absolute terms established at the global level by the United Nations for each indicator of SDG 3, provided that the region has not yet reached the level of progress established according to these targets.
2. Targets in absolute terms derived from the Sustainable Health Agenda for the Americas 2030 (ASSA 2030), for those indicators that had already exceeded the level of progress stipulated in the global SDG 3 targets at the time of analysis.

2.1 Targets taken from the SDG 3 definition for those indicators that only had a percentage reduction target without indicating an absolute baseline.

2.2 Relative targets derived from the ASSA 2030 for those indicators that did not have an agreed relative or absolute target when defining the goals linked to SDG3.

The Table 4 details the situation in relation to each indicator considered, as well as the origin of the target.

Table 4 - Progress indicators associated with SDG3 and their respective targets.

| ODS | Indicator detail | Target set with the United Nations | Agenda 2030 - Health for the Americas | Type of goal |
|------------|---|------------------------------------|---------------------------------------|--------------|
| SDG 3.1.1 | Maternal mortality rate | 70 | 30 | Absolute |
| SDG 3.1.2 | Deliveries attended by skilled health personnel | Not defined | Not defined | No goal |
| SDG 3.2.1 | Under-five mortality rate | 25 | 25 | Absolute |
| SDG 3.2.2 | Neonatal mortality rate | 12 | 12 | Absolute |
| SDG 3.3.1 | Number of new HIV infections | Not defined | -90% | Relative |
| SDG 3.3.2 | Incidence of tuberculosis | Not defined | -80% | Relative |
| SDG 3.3.3 | Estimated incidence of malaria | Not defined | -90% | Relative |
| SDG 3.3.4 | Hepatitis B surface antigen prevalence - children under 5 years of age. | Not defined | Not defined | No goal |
| SDG 3.3.5 | Number of people intervened against the ETD | Not defined | Not defined | No goal |
| SDG 3.4.1 | Age-standardized NCD mortality rate | 0,33 | -67% | Relative |
| SDG 3.4.2 | Age-standardized suicide rates | 0,33 | -67% | Relative |
| SDG 3.5.1 | Coverage of substance use disorder treatment interventions. | Not defined | Not defined | No goal |
| SDG 3.5.2 | Harmful use of alcohol (15+) 12-month prevalence | Not defined | -20% | Relative |
| SDG 3.6.1 | Estimated mortality rate due to traffic accidents | 0,5 | -50% | Relative |
| SDG 3.7.1 | Women with modern family planning methods | Not defined | 90 | Absolute |
| SDG 3.7.2 | Adolescent birth rate per 1,000 females | Not defined | -10% | Relative |
| SDG 3.8.1 | Coverage of essential health services | Not defined | 100 | Absolute |
| SDG 3.8.2 | Proportion of large household expenses 10%. | Not defined | Not defined | No goal |
| SDG 3.8.2 | Proportion of major household expenses 25% of total household expenses | Not defined | Not defined | No goal |
| SDG 3.9.1 | Mortality rate due to environmental and domestic air pollution | Not defined | Not defined | No goal |
| SDG 3.9.2 | Mortality rate attributed to unsafe WASH | Not defined | Not defined | No goal |
| SDG 3.9.3 | Mortality rate attributable to unintentional poisoning | Not defined | Not defined | No goal |
| SDG 3.a. 1 | Estimation of current prevalence of tobacco use | Not defined | -40% | Relative |
| SDG 3.b. 2 | Official development net medical research basic health care | Not defined | Not defined | No goal |
| SDG 3.b. 3 | Health centers with a basic set of essential medicines | Not defined | Not defined | No goal |
| SDG 3.c. 1 | Density of qualified health professionals | Not defined | Not defined | No goal |
| SDG 3.d. 1 | Basic capacity of the International Health Regulations | Not defined | Not defined | No goal |
| SDG 3.d. 2 | Percentage of bloodstream infections due to antimicrobial-resistant organisms | Not defined | -10% | Relative |

As for the analysis procedures for each of the 5 analytical strategies considered, the methodological steps adopted will be detailed, as well as the parameters considered for each analysis.

Analysis of the evolution of the average of the indicators at the regional level over time;

The analysis of the time trend and progress consisted of two stages. In the first stage for each subregion, the time series corresponding to the indicator was identified and in the second stage the progress achieved for the region was evaluated, taking 2015 as a reference.

- For each indicator with absolute targets, the current value was compared with the value established for the objective.
- In the case of indicators with relative targets, the expected percentage change was applied to the value shown in 2015 to obtain the reference values to be evaluated in 2030. Thus, for example, if for an indicator the relative target was a 90% reduction of the baseline value, the calculation of the expected value in 2030 was performed by applying a 90% reduction on the baseline value obtained in 2015.
- For indicators for which data were available, but for which there was neither a relative target nor an absolute target, a trend assessment was performed. Thus, to obtain a baseline figure in 2030, the average annual variation between the first year for which data was available and 2015 was analyzed. From there, it was possible to obtain a percentage change observed in the indicator prior to the baseline assessment.

A polarity categorization was then created for each indicator. For indicators with a positive polarity, it was established that the higher the indicator, the better the health situation of the country, such as the density of health professionals. For indicators with negative polarity, it was established that the lower the value attributed to the indicator, the better the health situation of the country, for example, mortality indicators.

For it to be possible to apply the percentage obtained through trend analysis to obtain baseline values for a target in 2030 to indicators that had no target identified, it was necessary for the polarity and the percentage obtained through trend analysis to be aligned. Thus, if the percentage obtained for the trend analysis of an indicator was negative and the polarity of the indicator was also negative, the percentage was applied to obtain the reference value in 2030; otherwise, the approach defined by the trend analysis could not be applied. Thus, whenever possible, the percentage obtained by trend analysis was applied to the value observed for 2015 in order to obtain a reference value for comparison in 2030. At the end of this process it was possible to obtain the following indicator ranking matrix:

Table 1 - Matrix of possibilities for progress on indicators

| | | Evaluation of trends | |
|------------------------|--------------------|--|-----------------------------------|
| | | Increase | Decrease |
| Evaluation of progress | At the finish line | Objective achieved and improving | Target achieved and deteriorating |
| | Off-target | Objective not achieved and in progress | Target not met and deteriorating |

In addition to the categorization of progress, an analysis of the average annual percentage change was performed. This analysis was based on the methodology developed by Sanhueza et

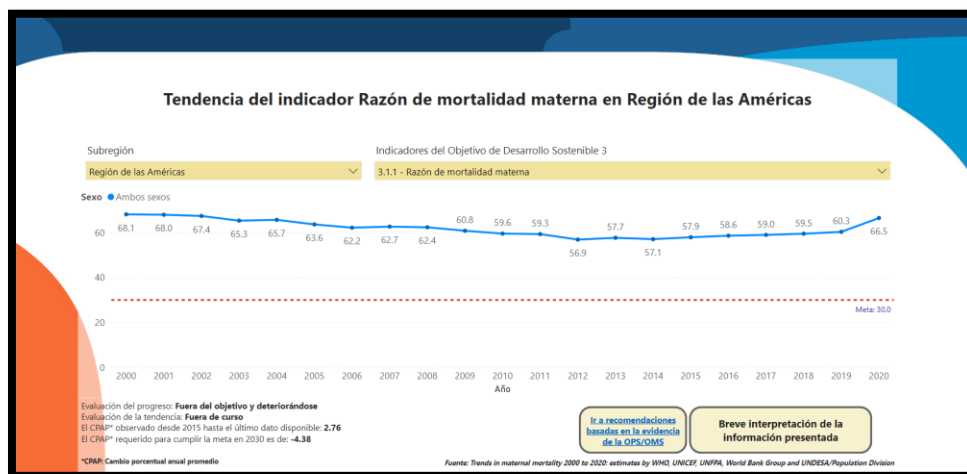
al 2020². If we have the value of a health indicator (SI) at two defined times, t₀ and t₁, it is possible to calculate its average annual percentage change (AAPC) using the following expression, with the use of natural logarithms.

$$CPAP = \frac{\ln(IS_{t_1}) - \ln(IS_{t_0})}{(t_1 - t_0)} \times 100$$

As a result, it presents the average percentage change of the indicator over a time interval. This calculation was applied twice, the first time to obtain the percentage change observed from 2015 to the latest available data and the second time to identify the percentage change required to achieve the 2030 targets. In the case of subregions that had already achieved the 2030 target, this analysis was not carried out because the necessary progress had already been made. In the case of indicators and subregions without sufficient data on progress, this analysis was also not carried out. The latter situation includes indicators and subregions without a minimum amount of data over time to allow comparison with the baseline year and indicators and subregions without possible targets.

Figure 3 shows in red lines the target to be achieved by 2030. Where there was a gender split in the data, a time series was prepared for each gender, as well as an analysis for both genders. Estimated progress, time trends and percentage change were calculated separately for each gender group. A Figure 3 exemplifies the results linked to the analyses cited in this section.

Figure 3 - Example of progress and time trend analysis.

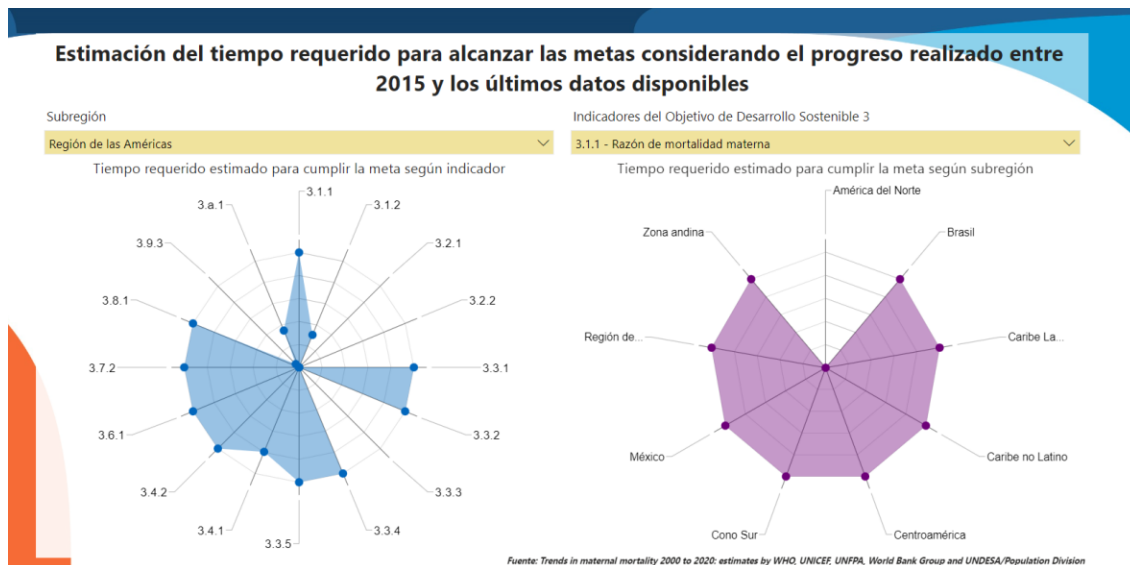


² Sanhueza A, Espinosa I, Mújica OJ, da Silva Jr JB. Leaving no one behind: a methodology for setting Sustainable Development Goal 3 health inequality reduction targets. Rev Panam Public Health 2020; [doi: 10.26633/rp.2020.155].

Estimated time required to reach the 2030 target

The analysis of the time elapsed to the achievement of the objective was based on the information obtained from the time series and the percentage changes identified between the reference year and the last year with available data. Thus, this analysis applied the average annual change identified to the reference values. Thus, the average annual percentage change between 2015 and the last year with available data was applied. In this way, it was possible to obtain an estimate of how much time each subregion would need to achieve the goals defined in the 2030 Agenda, taking into account the actual progress they have been showing in recent years. This analysis was carried out in two ways, one with grouping based on subregions, to identify the situation of each region with respect to all the indicators with data, and another based on indicators, to allow comparisons between the progress made by each subregion. Whenever the time frame for the targets exceeded 50 years, the analysis was interrupted, since the progress achieved during the years evaluated showed a trend toward zero. The results were presented in the form of a radar graph. The closer to the center of the graph, the greater the possibility of progress toward achieving the goals, as the axis running from the center to the extremes is arranged hierarchically with the years to achieve the goals starting at zero and ending at 50 years. The closer to the extremes, the longer it will take to reach the targets and therefore the less progress will be observed for the indicator or subregion in question. The following figures detail the time required to reach the targets described above.

Figure 4 - Time to target achievement by indicator



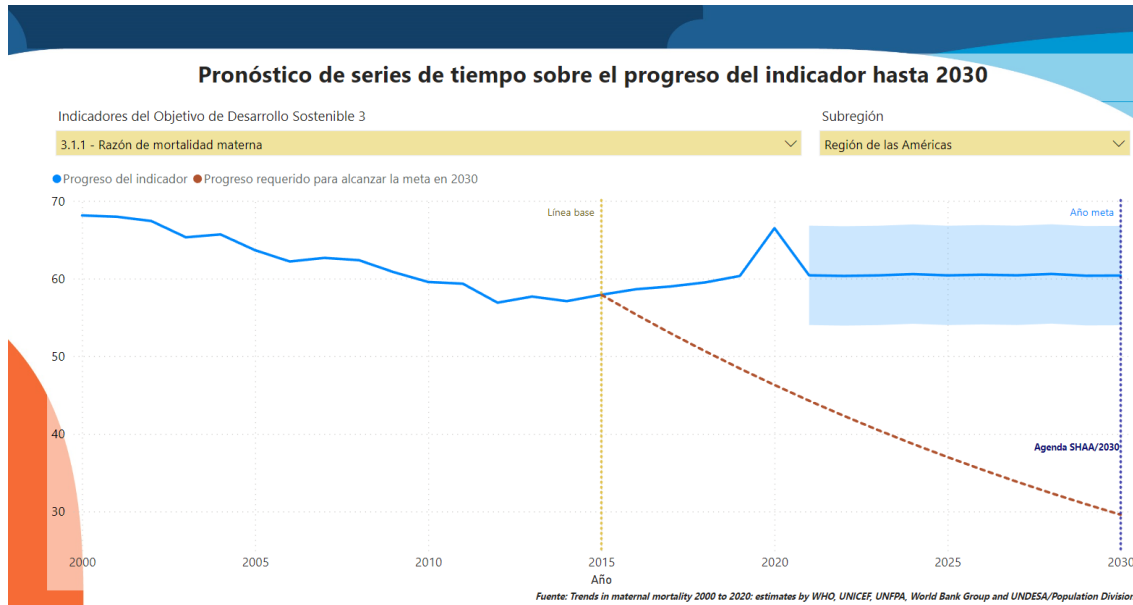
The forecast of the indicator's temporal progress.

The projection of the time series was performed using machine learning techniques. The analysis in question was based on projection principles using a training-testing approach. Using this approach, for each subregion and for each indicator, a projection model with the lowest possible error was fitted. Thus, for projection purposes, multiple models were fitted independently, always with the objective of obtaining the best possible results for the subregion-indicator dyad. The training-testing approach was carried out as follows:

1. Data for an indicator and a subregion are separated from the mass of data analyzed.
2. It is evaluated for how many points in time information is available and from that total the last 5 data points in time are excluded.
3. The remaining data points are used for the training phase. For this phase, different statistical time series projection models are applied to the data.
4. The results of each applied statistical model are compared with the actual values of the series and the difference between the actual values and the predicted values is evaluated.
5. The model with the smallest distance between actual and predicted, for the sequence of training points, is chosen as the most accurate for that indicator and subregion.
6. The most accurate model is used to predict the entire time series, including the 5 time points not previously used; this step characterizes the test phase.
7. The model is reevaluated and the one with the lowest level of error, considering as a metric the root mean square error (RMSE) statistic, is selected as the most appropriate model for the time series projection.
8. The confidence intervals of the prediction with the chosen model are calculated and the projection for the future of the time series sequence is made.
9. The year 2030 was defined as the limit for future projections.

The whole process of application, evaluation and choice of models is performed automatically, having as a reference the minimization of performance metric (RMSE). This metric was chosen because of its ability to handle different confounding factors, in terms of prediction accuracy for structured approaches according to regression logic. In total, five statistical models were tested: *Prophet*, *XGBoost*, *Random Forest*, *Linear Regression*, *Arima* and *Elasticnet*. This set of models was chosen to contemplate different approaches and assumptions, especially considering models with softer requirements in terms of data distribution shape and statistical validity assumptions. With the results of the predictions made, graphs were produced, taking as a reference the progress required to achieve the goals agreed in the 2030 agenda, as well as the values predicted using the previous approach, which Figure 5 exemplifies the graphic format used.

Figure 5 - Example of comparison of time series prediction with required progress.



The table with the accuracy metrics in relation to the models tested for each indicator can be found in Annex 1 of this document.

The evolution over time of social inequality between countries in the indicator

In order to assess the distribution of equity in relation to the health outcome obtained in the SDG3 indicators, analyses were performed based on two metrics: the absolute Kuznets gap and the relative Kuznets gap. This indicator measures the distance between groups of countries stratified by an equity parameter.

The absolute gap is a simple summary metric of health inequality and corresponds to the arithmetic difference in the value of the health indicator between two socially determined population groups, usually extreme groups of social position. This metric is calculated according to the following expression:

$$BA = ISq1 - ISq5$$

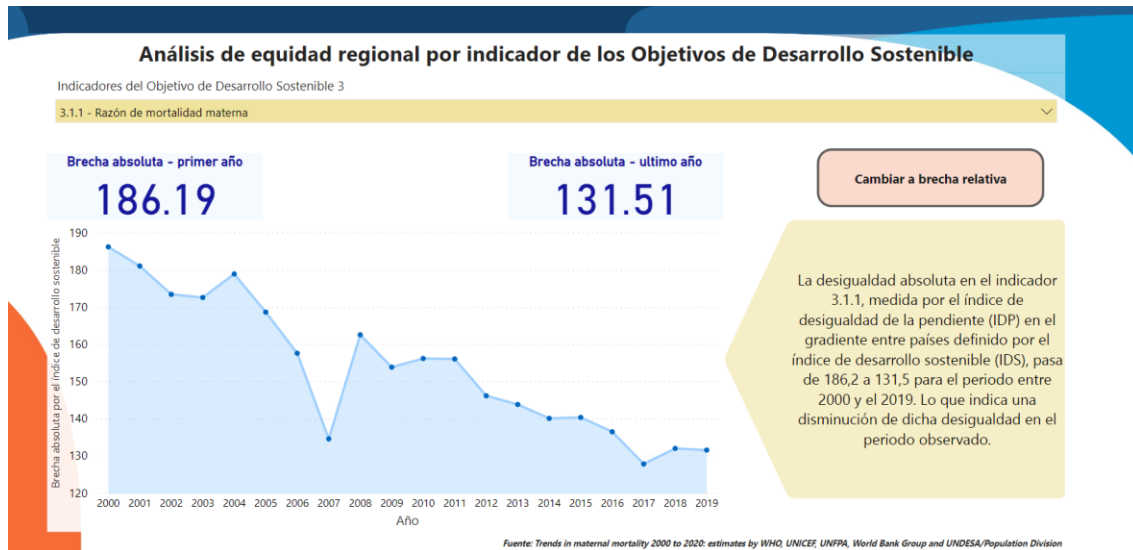
where *BA* is the absolute gap; *IS*, the health indicator; *q1*, the most disadvantaged social position quintile; and *q5*, the most advantaged social position quintile. The absolute gap is expressed in the same units of measurement as the health indicator; a zero (0) absolute gap value denotes the absence of inequality.

The relative gap is a simple summary metric of health inequality and corresponds to the arithmetic quotient of the value of the health indicator between two socially determined population groups, usually extreme groups of social position. This metric is calculated according to the following expression:

$$BR = \frac{ISq1}{ISq5}$$

where *BR* is the relative gap; *IS*, the health indicator; *q1*, the most disadvantaged social position quintile; and *q5*, the most advantaged social position quintile. The relative gap is expressed without units of measurement (i.e., its value represents the number of times the numerator is contained in the denominator). A relative gap value of one (1) denotes the absence of inequality. This parameter for this study was the Sustainable Development Index (SDIx). Based on this stratifier, the countries in each subregion were classified. Based on this hierarchy, a function was applied to calculate and extract the values associated with both gaps for each year for which information was available for the indicators. The following figure exemplifies the graphs developed to demonstrate the variations in relation to the equity metrics considered. The values associated with the absolute gap show the absolute distance for the indicator evaluated between the most and least advantaged groups in relation to the equity stratifier.

Figure 6 - Graphical demonstration of the absolute phase shift.



The magnitude of the changes observed in the averages and inequalities in the indicator

The last analytical step performed was a comparison between the levels of inequality identified with the absolute gradient metric. For each subregion, the absolute gradient indicator was calculated for 2015 and for the last year with available data. From these two values, we calculated the average annual change between the two points in time. The same average percentage change calculation was performed for each SDG3 indicator with data. The results for both metrics were plotted on a scatter plot to place the subregions into one of the following four possibilities:

- Quadrant 1: Subregional average improves and inequality among countries in the same subregion increases.
- Quadrant 2: The subregional average worsens and inequality among countries in the same subregion increases.
- Quadrant 3: Improves the subregional average and reduces inequality among countries in the same subregion.
- Quadrant 4: The subregional average worsens and inequality among countries in the same subregion decreases.

Successful progress toward the SDG 3 targets, "leaving no one behind," demands both average improvement and distributional improvement (Quadrant 3).

The Figure 7 details the graphical format used to demonstrate the results obtained from this analysis.

Figure 7 - Graphical detail of the comparison between inequalities and progress in the indicators associated with the SDGs.

