

JMP METHODOLOGY

2017 UPDATE & SDG BASELINES

March 2018



World Health
Organization

WHO
UNICEF



JMP

unicef



Contents

Summary	4
1. Identification of national datasets	6
2. Data extraction, cleaning and pre-processing	8
2.1 Data compilation	8
Year assignment	8
Geographic scope	8
Weighting	8
2.2 Technology classifications	8
Standard classifications: drinking water	8
Standard classifications: sanitation	9
Standard classifications: handwashing	9
2.3 Service level classifications: water	10
Within 30 minutes	10
On premises	10
Available when needed	10
Free from contamination	10
2.4 Service level classifications: sanitation	11
Not shared with other households	11
Containment	11
Emptying of on-site storage facilities	11
Treatment and disposal of excreta from on-site storage facilities	12
Wastewater transported to treatment	12
Wastewater treated	12
2.5 Data cleaning and adjustment	12
2.6 Data acceptance	12
3. Producing estimates	13
3.1 National, rural and urban estimates	13
3.2 Primary indicators	13
Special considerations	15
3.3 Ratios	15
Special considerations	16
3.4 Secondary indicators: drinking water	16
Data gaps	16
3.5 Secondary indicators: sanitation	17
Data gaps	17
3.6 Secondary indicators: hygiene	18
3.7 Annual Rates of Change	18
3.8 Wealth quintiles	19
3.9 Subnational	19
4. Country consultation	20

5. Calculation of regional and global estimates	21
5.1 Population data	21
5.2 Imputations for countries lacking data	21
Basic services	21
Safely managed services	22
Annex	
JMP Country Files	23

Summary

The WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP) was established in 1990 and has continuously refined the methods used for global monitoring. This document describes the methods and rules applied in the last JMP report, “Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines”¹. These methods supersede those used in previous JMP reports.

The JMP produces estimates for a total of 26 indicators related to water, sanitation and hygiene. Estimates are produced at the country, regional, and the global levels. All estimates are expressed as the proportion of the population which uses a specific type of water, sanitation, or hygiene technology or level of service.

Table 1: 26 indicators produced by the JMP

Water	The proportion of the population that uses...
W ₁	improved drinking water sources
W ₂	piped improved drinking water sources
W ₃	non-piped improved drinking water sources
W ₄	unimproved drinking water sources
W ₅	no drinking water facility (surface water)
W ₆	improved water sources exceeding 30 minutes collection time (limited drinking water services)
W ₇	improved water sources not exceeding 30 minutes collection time (basic drinking water services ²)
W ₈	improved water sources which are accessible on premises
W ₉	improved water sources which are available when needed
W ₁₀	improved water sources which are free from contamination
W ₁₁	safely managed drinking water services ³
Sanitation	The proportion of the population that uses...
S ₁	improved sanitation facilities
S ₂	improved sanitation facilities connected to sewers
S ₃	improved sanitation facilities connected to septic tanks
S ₄	improved pit latrines or other on-site improved facilities
S ₅	unimproved sanitation facilities
S ₆	no sanitation facility (open defecation ⁴)
S ₇	improved sanitation facilities which are shared (limited sanitation services)
S ₈	improved sanitation facilities which are not shared (basic sanitation services ²)
S ₉	sewer connections where wastes reach treatment plants and are treated
S ₁₀	on-site sanitation facilities where wastes reach treatment plants and are treated
S ₁₁	on-site sanitation facilities where wastes are disposed of in situ
S ₁₂	safely managed sanitation services ⁵
Hygiene	The proportion of the population that has...
H ₁	a handwashing facility on premises
H ₂	a handwashing facility on premises with soap and water available (basic handwashing facility ^{2,5})
H ₃	a handwashing facility on premises lacking soap and/or water (limited handwashing facility)

1 <https://washdata.org/report/jmp-2017-report-final>

2 Indicator for SDG target 1.4

3 Indicator for SDG target 6.1

4 Additional indicator for SDG target 6.2

5 Indicator for SDG target 6.2

Wherever possible, these estimates are disaggregated by urban and rural location. The JMP is furthermore developing a database of estimates which can also be disaggregated by wealth quintile, sub-national region and other relevant stratifiers of inequality. Many of the datasets used for producing estimates are household surveys and censuses which collect water, sanitation and hygiene information at the household level. Such data cannot be disaggregated to provide information on intra-household variability, e.g. differential use of services by gender, age, or disability. The JMP seeks to highlight individual datasets which do allow assessment of intra-household variability, but these are not numerous enough to integrate into the main indicators estimated in JMP reports.

The production of estimates follows a consistent series of steps:

- 1. Identification of appropriate national datasets**
- 2. Extraction of data from national datasets into harmonized tables of data inputs**
- 3. Use of the data inputs to model country estimates**
- 4. Consultation with countries to review the estimates**
- 5. Aggregation of country estimates to create regional and global estimates**

This document describes each of these steps in detail.

Changes introduced in the 2017 Update

The statistical methods used in the 2017 Update differed from those used in the 2015 update and previous reports in a number of ways:

- Most obviously, the 2017 Update introduced the new indicators of **safely managed drinking water services** and **safely managed sanitation services**, and developed new methodologies for calculating these indicators.
- The 2017 Update further developed the methods used to estimate the **handwashing with soap and water** indicator. In the 2015 Update, estimates of handwashing for a country were taken from a single household survey, and limited to that year. In the 2017 Update, multiple surveys were used (where available) and standard JMP regression rules were used to produce estimates for multiple years.
- New definitions of **basic and limited services** were introduced, which build on the “use of improved facilities” indicators used in the MDG period. While the JMP method continues to produce estimates of “use of improved facilities”, the bulk of the analysis is made on the new “basic services” indicators.
 - ◆ **Basic drinking water services** are defined as drinking water from an improved source, provided collection time is not more than 30 minutes for a roundtrip including queuing.
 - ◆ **Limited drinking water services** are defined as drinking water from an improved source, where collection time exceeds 30 minutes for a roundtrip including queuing.
 - ◆ **Improved water sources now include** packaged and delivered water, considering that both can potentially deliver safe water.
 - ◆ **Basic sanitation services** are defined as use of improved sanitation facilities which are not shared with other households. This is identical to the “improved but not shared” category used in previous reports.
 - ◆ **Limited sanitation services** are defined as use of improved sanitation facilities which are shared with other households. Previous reports used the term “shared but otherwise improved sanitation facilities”.
 - ◆ **Improved sanitation** in the 2017 Update refers only to the type of facility used, irrespective of whether the facilities are shared by more than one household. In the 2015 Update and earlier reports, improved sanitation was defined to exclude shared facilities.
- Many new data sets were added for production of the 2017 Update. However, statistical analysis was for the most part limited to data sets from 2000 or later. Accordingly, estimates in the 2017 Update were only produced for 2000-2015, where data permit. The JMP no longer produces estimates for years prior to 2000.
- As for all JMP progress updates, estimates in the most recent 2017 report supersede estimates from all previous reports.

1. Identification of national datasets

The JMP has been reporting on water sanitation and hygiene since its inception in 1990 and has amassed a large database. Every 1-2 years, the JMP reviews available national datasets, updates the global database, and publishes updated estimates. The data search is largely done by JMP staff members, systematically visiting the websites of national statistical offices, and key sector institutions such as ministries of water and sanitation, regulators of drinking water and sanitation services, etc. Other regional and global databases are also reviewed for new datasets. UNICEF and WHO regional and country offices are asked to help identify newly available household surveys, censuses and administrative datasets, including those submitted via UNICEF's internal CRING system (Country Reporting on Indicators and Goals).

During the Millennium Development Goal (MDG) period, the JMP database included datasets going back to 1990 and even earlier. For monitoring the Sustainable Development Goals (SDGs), the JMP will restrict

analysis to datasets from 2000 and later, with a few exceptions in cases where it would be impossible to generate an estimate without using an earlier dataset from 1990-2000. For the 2017 Update, pre-2000 data were drawn upon from for three countries, areas or territories.

Datasets included in the JMP database include:

- Censuses, which in principle collect basic data from all people living within a country. Censuses are always led by national statistical offices.
- Household surveys, which collect data from a subset of households. These may target national, rural, or urban populations, or more limited project or sub-national areas. An appropriate sample design is necessary for survey results to be representative, and surveys are often led by or reviewed and approved by national statistical organizations.
- Administrative data, may consist of information collected by government or non-government entities involved in the delivery or oversight of services. Examples include

Table 2: Datasets available and used in the JMP 2017 update

Year	Censuses		Household Surveys		Administrative datasets		Other		Total	
	used	total	used	Total	used	total	used	total	Used	Total
2000	27	29	88	98	41	54	30	54	186	235
2001	29	33	41	51	47	58	3	16	120	158
2002	15	16	59	68	50	59	4	20	128	163
2003	5	6	99	118	42	52	5	21	151	197
2004	6	7	53	65	51	63	5	25	115	160
2005	8	9	72	80	90	106	5	31	175	226
2006	14	18	95	101	71	86	6	28	186	233
2007	9	11	75	83	79	94	7	35	170	223
2008	13	14	68	77	75	81	6	32	162	204
2009	10	12	68	74	85	112	6	31	169	229
2010	38	40	75	82	93	101	6	31	212	254
2011	35	35	74	79	111	129	6	36	226	279
2012	14	14	84	90	160	191	10	39	268	334
2013	6	6	73	81	128	158	15	46	222	291
2014	7	8	83	92	104	118	5	31	199	249
2015	4	6	46	55	81	88	8	23	139	172
2016	1	1	23	25	19	20	0	0	42	46
TOTAL	241	265	1180	1323	1329	1572	127	499	2876	3658

water and sanitation inventories and databases, and reports of regulators.

- Other datasets may be available such as compilations by international or regional initiatives (e.g. IB-NET), studies conducted by research institutes, or technical advice received during country consultations.

Some datasets reviewed by the JMP are not representative of national, rural or urban populations, or may be representative of only a subset of these

populations (e.g. the population using piped water supplies or sewer connections). The JMP will include datasets in its database when they represent at least 20% of the population of interest. However, not all datasets in the database are used in the production of estimates (see section 2.6, Data Acceptance).

The JMP database used to produce the 2017 update included 3,653 datasets, of which 2,871 were used to generate estimates (*see Table 2*).

2.

Data extraction, cleaning and pre-processing

Once datasets are identified, national data on water, sanitation and hygiene are extracted from the original sources and recorded in standardized spreadsheets. Data from each country, area, or territory are recorded in individual spreadsheets (the JMP country files), with water, sanitation, and hygiene data recorded on separate sheets (see Annex 1 for details). The JMP maintains individual country files for each of the 232 countries, areas, and territories which are included in the UN Population Division's World Population Prospects tables (2015 revision). However, three of these entities (Caribbean Netherlands, Holy See, and Western Sahara) have no available data, and their country files are blank. Country files can be downloaded from the JMP website: www.washdata.org

2.1 Data compilation

The JMP compiles national data from a wide range of different data sources. Where possible, the JMP uses original microdata to produce its own tabulations, which are then transferred to the country files. However, in many cases microdata are not readily accessible so relevant data are transcribed from reports available in different formats (PDFs, Word files, Excel spreadsheets, etc...).

Year assignment

Datasets are assigned a single year in the JMP database, even when data collection and publication span several years. The year assigned to a dataset is the last year in which data collection was made. Thus a survey conducted from 1 October 2010 through 2 January 2011 would be assigned to 2011.

Some panel surveys collect multiple rounds within a single year, or disaggregate findings temporally (e.g. by rainy and dry season). In this case, where possible, data will be combined into a single set of statistics representing average values for the year.

Geographic scope

The JMP country files include data that are representative of national populations, as well as data that are representative of urban and rural populations. Where possible, data for all three domains are extracted but in some cases data are either unavailable or it is not

possible to disaggregate them for urban and rural areas. Data that are not representative of any of these domains are generally not extracted into the country files (or are inserted into the country files but flagged as non-representative and not used for generating estimates). The JMP is currently developing methods for extracting data which are representative of sub-national regions and this is expected to become a standard feature in coming years.

Weighting

In order to produce statistics about the proportion of the *population* using a particular service, rather than the proportion of *households* having that service, the JMP uses population weights for household survey and census data. Where microdata do not include population weights, but do include household weights and the number of de jure household residents, these are multiplied together to generate population weights which are subsequently used in tabulations. Where microdata do not include population weights, or where microdata are not available, the JMP will use statistics that are based on household weighting. The JMP in most cases will not use data that do not have either household or population weights, or where data are not tabulated for the proportion of the population, or household/dwelling.

2.2 Technology classifications

The JMP uses a standard classification system to categorise populations using different water, sanitation and hygiene technologies. Data extracted from datasets are matched to the corresponding standard JMP classifications. The JMP gives priority to data coming from population-based sources such as censuses and household surveys which collect information directly from household members.

Standard classifications: drinking water

Drinking water sources are classified into two hierarchical levels based on technology. Where classification to the second level is not possible, classification is made to the first level only. Each second-level classification is considered as either an "improved" or an "unimproved" drinking water source. The use of "improved" drinking water sources served as the main MDG indicator on

drinking water, and forms the foundation of the new SDG indicators on drinking water.

Table 3: Classification of drinking water technologies

First level classification	Second level classification	Improved	Unimproved
Tap water	Piped water into dwelling	X	
	Piped water to yard/plot	X	
	Public tap, standpipe	X	
	Other	X	
Ground water	Tubewell, borehole	X	
	Protected well	X	
	Protected spring	X	
	Unprotected well		X
	Unprotected spring		X
Rainwater	Covered cistern/tank	X	
	Uncovered cistern/tank	X	
Packaged water*	Bottled water	X	
	Sachet water	X	
Delivered water*	Cart with small tank/drum	X	
	Tanker truck provided	X	
Surface water	River		X
	Lake		X
	Dam		X
	Pond		X
	Stream		X
	Irrigation channel		X
Other	Other improved	X	
	Other unimproved		X
DK/Missing			X

*Packaged and delivered water were previously categorised as unimproved for MDG monitoring

Standard classifications: sanitation

Sanitation facilities are classified into two hierarchical levels based on technology. Where classification to the second level is not possible, classification is made to the first level. Each second-level classification is considered as either an “improved” or an “unimproved” type of sanitation facility.

Standard classifications: handwashing

The new global SDG indicator for handwashing is the proportion of population with handwashing facilities with soap and water at home. Handwashing facilities can consist of a sink with tap water, but can also include other devices that contain, transport or regulate the flow of water. Buckets with taps, tippy-taps and portable basins are all examples of handwashing facilities. Bar soap, liquid soap, powder detergent and soapy water all count as soap for monitoring purposes.

Table 4: Classification of sanitation technologies

First level classification	Second level classification	Improved	Unimproved
Flush toilets	to piped sewer system	X	
	to septic tank	X	
	to pit	X	
	to unknown place/not sure/DK	X	
	to open drain		X
	to elsewhere		X
Pour flush latrines	to piped sewer system	X	
	to septic tank	X	
	to pit	X	
	to unknown place/not sure/DK	X	
	to elsewhere		X
Dry latrines	Ventilated Improved Pit latrine	X	
	Composting toilets	X	
	Pit latrine with slab	X	
	Pit latrine without slab/open pit		X
	Hanging toilet/hanging latrine		X
	Bucket latrine		X
No facility	Bush, field		X
Other	Other improved	X	
	Other unimproved		X
DK/Missing			X

People living in households that have a handwashing facility with soap and water available on premises are classified as having basic facilities. Households that have a handwashing facility but lack water and/or soap are classified as having limited facilities. In some cultures, ash, soil, sand or other materials are used as handwashing agents, but these are less effective than soap and are therefore counted as limited handwashing facilities.

The JMP uses data from household surveys in which the enumerator observes the handwashing facility and confirms the presence or absence of soap and water. Datasets that include self-reported availability of handwashing facilities, soap and water may be included in the JMP database and country files, but in most cases are not used for making estimates.

In some parts of the world (notably Central America) households sometimes do not give permission for survey enumerators to enter the premises and observe handwashing facilities. These households are excluded from calculations of the proportion of households having handwashing facilities.

2.3 Service level classifications: water

Additional datasets are used to calculate the proportion of households that use improved drinking water sources which meet different service level criteria. These proportions are expressed as ratios (see Section 3.3). The service level criteria derive from the human right to safe drinking water and include accessibility (within 30 minutes, on premises), availability when needed, and quality. Service level information may come either through population-based or administrative datasets or other national data sources.

Within 30 minutes

Many household surveys and some censuses collect data on the amount of time that is typically required to collect water. Ideally, the survey question should specify the amount of time that is required for each round trip (i.e. to get from the dwelling to the water collection point, queue for water, fill containers and return to the dwelling). Time spent socializing (outside of queuing) should not be included in the total number of minutes.

Most surveys record the number of minutes reported as a number, but others record ranges (e.g. 0-30 minutes). When a number is recorded, households using improved sources and reporting collection times ranging from 0 to 30 minutes (inclusive) are classified as having a drinking water source within 30 minutes. When ranges are reported, all ranges having an upper limit of 30 minutes or less are counted.

If surveys record the time for a one-way trip to a water source, this number is multiplied by two. If one-way travel time is recorded as a range, the JMP uses ranges with an upper limit of 15 minutes.

In countries with no information on time to collect but data on the proportion of the population with a water source within a maximum distance (e.g. 250 meters) this has been used as a proxy.

On premises

A water supply is considered as being “on premises” when water is collected from a point located within the dwelling, plot or yard. Piped water is commonly available on premises, but other water supply types may also be located on premises.

The JMP classifies households as having water supplies on premises when a survey collects information on the location of the water point, either as part of the response category or as a standalone question. Additionally, in

surveys that collect data on collection times, households reporting a collection time of zero minutes are classified as having water on premises.

In countries where the majority of the population uses piped water supplies, data on location and collection time are often not available. In such cases all piped water is considered to be located on premises, as well as 50% of non-piped improved water supplies.

Available when needed

The human right to safe drinking water specifies that water should be “available continuously and in a sufficient quantity to meet the requirements of drinking and personal hygiene, as well as of further personal and domestic uses, such as cooking and food preparation, dish and laundry washing and cleaning. [...] Supply needs to be continuous enough to allow for the collection of sufficient amounts to satisfy all needs, without compromising the quality of water.”⁶

“Availability of water when needed” is complex and is dependent on the quantity of water needed, the quantity of water available, the continuity and reliability of supplies, the capacity for in-home storage. Where available, the JMP uses data from population-based sources and a variety of questions have been included in national household surveys to date⁷. In cases where household respondents either report having access to sufficient water when needed, or having water available at least 50% of the time (i.e. at least 12 hours per day or 4 days per week) then the JMP classifies them as having drinking water services which are available when needed.

In the absence of data from population-based sources, data from administrative sources have also been used. Piped water systems that provide water for at least 12 hours per day or 4 days per week are categorised by the JMP as “available when needed”, although it is recognized that this may fall short of full realization of the human right.

Free from contamination

For global monitoring purposes “free from contamination” requires that drinking water should meet microbiological and priority chemical water quality standards. The microbiological standard applied is that no *E. coli* should be detected in a 100 mL sample. Thermotolerant coliform counts are another commonly used faecal indicator, and are considered as an acceptable alternative to *E. coli*⁸. Total coliforms are not considered as suitable indicators of faecal contamination, but can be used as indicators of treatment efficiency. If countries have data on total coliform levels in drinking water but not *E. coli* or

6 United Nations General Assembly, Report of the independent expert on the issue of human rights obligations related to access to safe drinking water and sanitation, Catarina de Albuquerque on 1 July 2010: A/HRC/15/31/Add.1', <<http://daccess-ods.un.org/access.nsf/Get?Open&DS=A/HRC/15/31/Add.1&Lang=E>>

7 For examples see JMP Thematic Report on Safely Managed Drinking Water Services <<https://washdata.org/report/jmp-2017-tr-smdw>>

8 Guidelines for Drinking Water Quality (4th edition), WHO, Geneva, 2011.

thermotolerant coliform data, the JMP uses compliance with total coliforms, although this is a more stringent indicator. Data on compliance with residual chlorine standards, while valuable for operational monitoring and surveillance, are not considered by the JMP as demonstrating compliance with microbiological standards.

At the global level, the priority chemical contaminants are arsenic and fluoride. The JMP collects data on compliance with relevant national standards, and where possible uses compliance with the WHO guideline values (10 µg/L and 1.5 mg/L, respectively).

Ideally the points at which water supplies are tested for compliance should be the location where users collect water, however in some cases water quality samples are collected at other locations, for example within the distribution network of a piped supply. The JMP prefers to use water quality surveillance data that are collected regularly throughout the year and cover the entire population of interest. A one-time collection of data from a representative sample of the population, such as water quality testing through household surveys, can also be used, recognising that such data only provide a "snapshot" in time and may therefore under-estimate the proportion of supplies that are contaminated.

Frequently water quality data are available only for formal supplies, or only for piped supplies. Such data are applied towards the entire population using improved supplies (in urban and/or rural areas) as long as the population to which the data relate is at least 80% of the population of interest (e.g. the population reporting use of piped water).

In countries where information is available for microbiological contamination as well as for arsenic and/or fluoride, the microbiological and chemical data may come from different sources. Therefore, it may be possible to state the percentage of population with supplies that are free from *E. coli*, or free from arsenic, or free from fluoride, but it may not be possible to state the percentage of population with supplies that are free from *E. coli*, arsenic and fluoride. To ensure comparability in such cases, the minimum value of the available parameters (i.e. supplies free from *E. coli*, or arsenic, or fluoride) will be used. However, no estimate will be made in the absence of data on microbiological compliance.

2.4 Service level classifications: sanitation

The definition of "safely managed sanitation services" is informed by the excreta flow diagram concept, in which faecal wastes from different types of sanitation facilities are tracked through stages of containment, emptying, transport, treatment, and reuse or final disposal (Fig. 1).

These stages are followed separately for wastes flushed into sewer networks, and for wastes stored in on-site facilities such as septic tanks and pit latrines⁹.

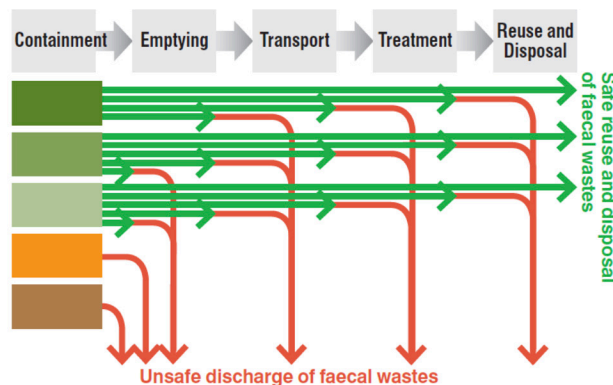


Figure 1: Excreta flow diagram¹⁰

Not shared with other households

Many household surveys ask respondents if the sanitation facility used by members of the household is shared with other households. Public toilets, as well as privately-owned sanitation facilities which are shared by two or more families, are classified as shared facilities.

Containment

Containment refers to the extent to which the sanitation facilities which people use ensure that excreta are safely separated from human contact at the household and community level. In cases where data exist on the proportion of people using improved sanitation systems which do not effectively contain faeces, (e.g. sewer lines, latrines or septic tanks that leak faeces into the environment) the JMP adjusts estimates of safely managed sanitation accordingly.

Emptying of on-site storage facilities

The JMP seeks data on the proportion of people using on-site sanitation storage facilities (septic tanks or latrines) which have ever been emptied. These data may come either from population-based data sources such as household surveys, or from administrative records. Information on the time since most recent emptying, the way excreta were emptied, or the use of protective equipment to prevent occupational exposures during emptying may be collected from such datasets but have not been used in calculating estimates.

Treatment and disposal of excreta from on-site storage facilities

This refers to the proportion of excreta emptied from on-site storage facilities (septic tanks or latrine pits) which are either buried on premises or transported, usually by cart, truck or tanker, and delivered to treatment plants

⁹ For further information see forthcoming JMP Thematic Report on Safely Managed Sanitation Services and Hygiene

¹⁰ <https://washdata.org/report/jmp-2015-report>

(regardless of the type of treatment plant) or discharged into sewer networks. Some household surveys and administrative sources collect information on the frequency and method of emptying on-site sanitation systems and whether excreta are delivered to treatment or unsafely discharged.

Excreta from on-site sanitation facilities may be transported to wastewater treatment plants or to specially designed faecal sludge treatment plants. Excreta delivered to wastewater treatment plants providing at least secondary treatment are classified as safely managed. Excreta delivered to faecal sludge treatment plants are classified as safely managed if both the liquid and solid fractions are treated. For the 2017 Update, data on treatment of excreta removed from on-site sanitation facilities were only available for two countries: Japan and Norway.

Wastewater transported to treatment

This refers to the proportion of excreta flushed into sewer systems which are transported along with wastewater and delivered to treatment plants (regardless of the type of treatment plant). Wastewater may not be transported to treatment due to exfiltration, pump failure, breaks or blockages in the sewer network system, or may discharge directly to open drains, water bodies or open ground. Data on the proportion of wastewater which is lost in transportation are rare.

Wastewater treated

This refers to the proportion of wastewater reaching wastewater treatment plants which receives at least secondary (biological) treatment. Wastewater which receives primary treatment is not considered safely managed, unless the effluent is discharged in a way that precludes further human contact (e.g. through a long ocean outfall).

If data are available for conventional classes (primary, secondary, tertiary, advanced) as well as for ambiguous categories (e.g. "other"), the ambiguous categories are generally not considered as safely managed. Where the only available information on treatment is ambiguous (e.g. "treated") the JMP assumes at least secondary treatment but seeks clarification during country consultations.

2.5 Data cleaning and adjustment

Where microdata are available, JMP tabulations are compared against any tabulations reported in e.g. survey reports or administrative reports. In case of inconsistencies, the questionnaire categories can be checked to identify possible errors.

Where response categories are ambiguous or do not map clearly into JMP standardized categories, adjustments can be made drawing upon data from other datasets for the same country and time period. For example, some household surveys record data on the population using ambiguous classifications such as "traditional wells".

Some traditional wells, but not all, meet the criteria for improved sources, so this population cannot immediately be classified as using an improved or unimproved water source. If another dataset is available which distinguishes between unambiguous classes (e.g. the proportion of all wells which are improved), this ratio can be applied to the survey with the ambiguous classification. If there are multiple datasets from which ratios can be generated, the dataset closest in time to the ambiguous data is used.

In a small number of cases, all datasets within a country use ambiguous classifications. In this case, there should be a specific consultation with the country, and in the absence of other information 50% of the ambiguous technologies can be considered as "improved", and 50% "unimproved". These adjustments are made in the JMP country files and highlighted in the notes.

2.6 Data acceptance

In some cases data that have been extracted and shown in the JMP country files are excluded from use in further analysis. Some reasons that can cause a dataset to be excluded include:

- Communication from national authorities that the data are not considered reliable or appropriate for use
- The classification of the data is based on few generic categories which are not aligned with JMP categories
- The "other" category is very large
- Data were available for the national population but not for the urban and rural ones (unless analysis for that country was done only at national level)
- Data were not representative of the national, urban or rural populations
- Data were representative of less than 80% of the national, urban or rural populations
- Sums of categories which should be mutually exclusive and exhaustive are far from 100%
- Data are markedly different from multiple other data points from a similar timeframe

For any of these reasons, a dataset may be excluded from calculation of any of the indicators reported by the JMP. In some cases a dataset can be used for one or more but not all indicators, because of variable data availability and quality. For example, a household survey might yield reliable data on "improved sanitation" but unreliable data distinguishing sewer connections from on-site sanitation systems, because of ambiguous question wording or inadequate training of survey teams.

3. Producing estimates

3.1 National, rural and urban estimates

Separate estimates are made for urban and rural areas, and national estimates are generated as weighted averages of the two, using population data from the most recent report of the United Nations Population Division. For the 2017 JMP Update, the proportion of the population living in urban areas was taken from the World Urbanization Prospects 2014 revision¹¹.

In 42 out of 232 countries, areas and territories, data inputs were not available disaggregated into urban and rural populations, and national estimates were generated directly. In eight countries, areas or territories, the entire population is classified as either urban or rural. In these cases national estimates are identical to the urban or rural estimates.

The safely managed drinking water and sanitation estimates W_{11} and S_{12} require combining multiple elements. In some cases a critical element is only available at the national level. This is often the case for wastewater treatment, or drinking water quality. In these cases only national estimates for the indicators are produced.

3.2 Primary indicators

The JMP estimates nine primary indicators directly from data inputs, all of which refer to the proportion of the population using or having access to water, sanitation, or hygiene facilities (see Table 5). These indicators are estimated using a set of “JMP estimation” rules for interpolation, extrapolation, and extension of estimates.

Interpolation rules (within the years covered by data inputs)

1. If data points are at least five years apart, estimates are interpolated using ordinary least squares linear regression.
2. If data points are less than five years apart, estimates are interpolated using a simple average.

Table 5: Nine primary indicators

Water	The proportion of the population that uses...
W_1	improved drinking water sources
W_2	piped water drinking water sources
W_5	no drinking water facility (surface water)
Sanitation	The proportion of the population that uses...
S_1	improved sanitation facilities
S_2	Improved sanitation facilities connected to sewers
S_3	Improved sanitation facilities connected to septic tanks
S_6	no sanitation facilities (open defecation)
Hygiene	The proportion of the population that has...
H_1	a handwashing facility on premises
H_2	a handwashing facility on premises with soap and water available (basic handwashing facility)

Extrapolation rules (outside the years covered by data inputs)

3. If there are at least two data points, estimates are extrapolated using ordinary least squares linear regression (or simple average, if data points are less than five years apart) forwards two years from the latest data point, and backwards two years from the earliest data point.
4. If extrapolation results in estimates above 100% or below 0%, the estimates are fixed at 100% and 0%, respectively.
5. If there is only one data point, no extrapolation is made.

Extension rules (beyond the years covered by data inputs and extrapolation)

6. Estimates are extended forwards four years from the latest estimate, and backwards four years from the earliest estimate
7. Estimates are extended forwards or backwards indefinitely, if the estimate is 99.5% or greater, or 0.5% or less.

Estimates are made for all years possible within the period from 2000 to two years prior to the year of reporting (the 2017 update included estimates for all years between 2000 and 2015).

¹¹ POP/DB/WUP/Rev.2014/1/F02, accessed June 18, 2017 from <https://esa.un.org/unpd/wup/>

In making primary estimates, data inputs from population-based sources (household surveys, censuses) or administrative sources (e.g. data from ministries, regulators) may be used. Where both exist for the same population the JMP will generally give preference to population-based sources. In either case data will only be used if they cover at least 80% of the relevant population (national, urban, or rural).

Examples

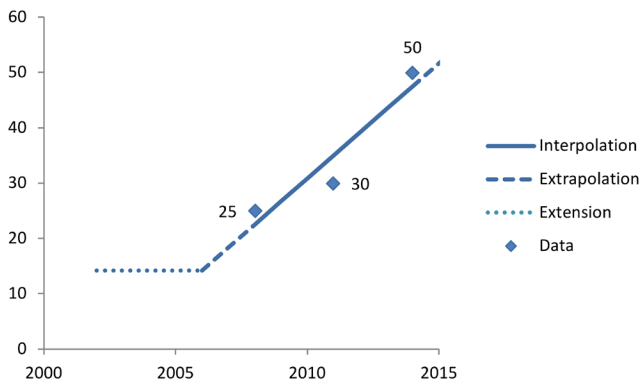


Figure 2: Three data points are available, spanning at least five years. Interpolation is made from 2008-2014 with linear regression. Extrapolation is made backwards for two years, from 2008 to 2006, and this estimate is extended backwards for four years, until 2002. Extrapolation is made forwards from 2014 to 2015, but not further, for the 2017 Update.

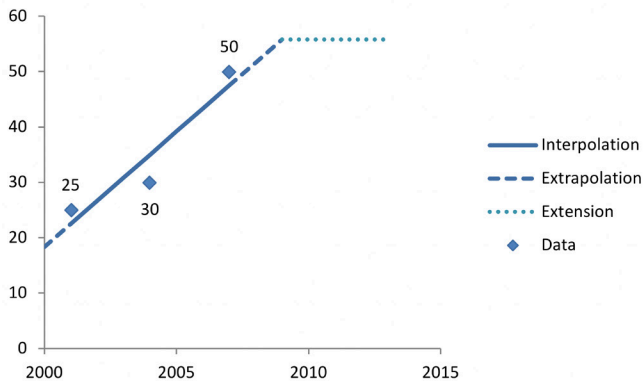


Figure 3: Three data points are available, spanning at least five years. Interpolation is made from 2001-2007 with linear regression. Extrapolation is made backwards from 2001 to 2000, and forwards from 2007 to 2009. The 2009 estimate is extended for four years to 2013, and no estimate is made for 2014 or 2015.

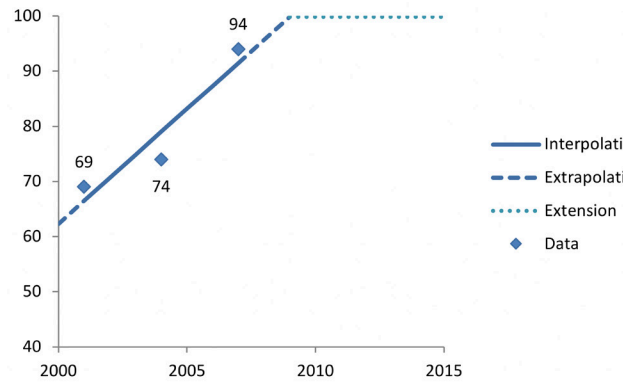


Figure 4: Three data points are available, spanning at least five years. Interpolation is made from 2001-2007 with linear regression. Extrapolation is made backwards from 2001 to 2000, and forwards from 2007 to 2009. The 2009 estimate of 99.8 % is extended to 2015, since it is above 99.5%.

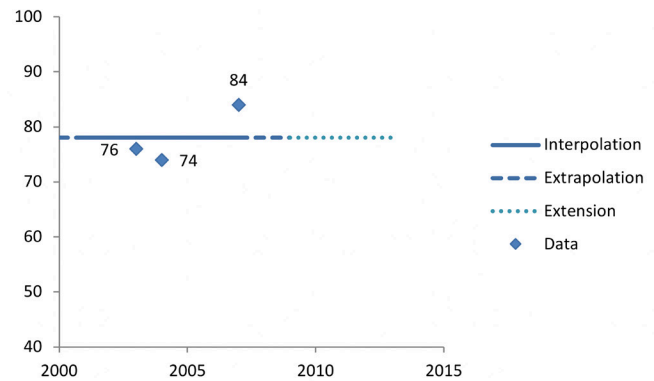


Figure 5: Three data points are available, spanning less than five years. Interpolation is made from 2004-2007 by averaging. Extrapolation of this averaged value is made backwards from 2003 to 2001, and forward from 2007 to 2009. The 2001 estimate is extended backwards to 2000, and the 2009 estimate is extended forward to 2013. No estimate is made for 2014 or 2015.

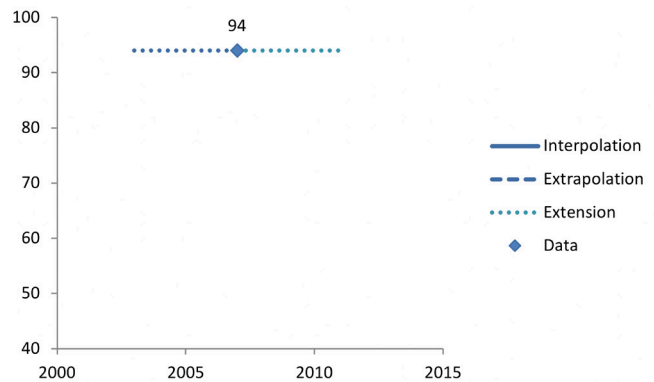


Figure 6: Only one data point is available. No interpolation or extrapolation is made. The 2007 estimate is extended backwards to 2003, and forward to 2011. No estimate is made for earlier or later years.

One implication of the JMP estimation rules is that when a new data point is added, it can change the estimates for all years, even the 'baseline' year of 2000. Figure 7 shows the impact of adding an additional data point to the dataset shown in Figure 3. When only three data points are available, the grey line shows an estimate of 56 in 2013, and because the most recent data point is from 2007 no estimates can be made for 2014 or 2015. If a fourth data point from 2010 is added, the slope of the line changes, and estimates can be made for both 2014 and 2015. Accordingly, the estimates published in each JMP update report supersede the estimates published in previous update reports.

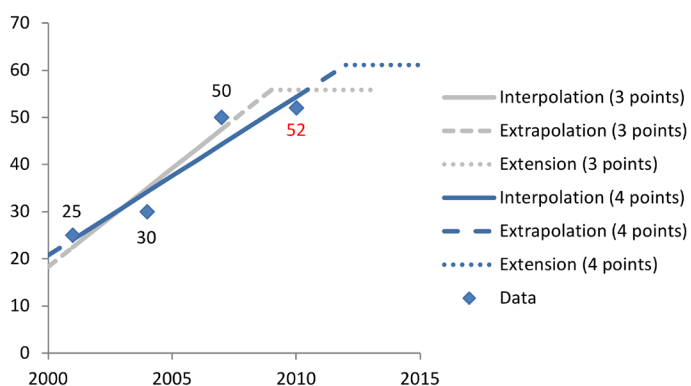


Figure 7: The impact of adding additional data points to JMP estimates.

Evaluation of alternatives to linear regression

In 2015 the JMP convened an Expert Group Meeting to review the methods used for interpolation, extrapolation and extension. Alternatives to linear regression were considered including logistic regression, piecewise spline, and loess. While the datasets in the JMP database show evidence of some nonlinear trends for some parameters over time, the nonlinear models assessed require large numbers of datapoints to allow interpolation of estimates that are superior to simple linear regression. For many countries, and especially for the service level indicators, there are not enough datapoints to justify the use of nonlinear methods. Nonlinear techniques are also limited in their ability to extrapolate even a few years, which is often necessary for the JMP's estimation method. Accordingly, the JMP has continued to use linear regression, with the set of rules regarding interpolation, extrapolation and extension described above¹².

Special considerations

Piped water is a subset of improved water, so piped water (W_2) should never be greater than improved water (W_1). Likewise sewer systems and septic tanks are both subsets of improved sanitation, so the sum of S_2 and S_3 should never exceed S_1 . Because separate regressions are made for each of these indicators, and may draw on different

sets of data, it is possible for these unrealistic situations to be estimated, in which case the JMP applies standard corrections. If the estimate for piped water exceeds the estimate for improved water, W_2 is set as equal to W_1 . If the sum of the estimates for sewer connections and septic tanks exceeds the estimate for improved sanitation, S_2 and S_3 are both proportionally reduced so that their sum is equal to S_1 , and improved latrines (S_4) is set to zero.

Estimates for these subsets of improved water and sanitation (W_2, S_2, S_3) are extended to cover all years where estimates are made for improved water W_1 or sanitation S_1 , respectively.

In some countries estimates for sewer connections and septic tanks are only available at the national level while improved sanitation can be calculated at rural and urban areas. In these cases a weighted average is used for the national estimate of improved (not shared) sanitation and this is split into sewer, septic and improved latrines and other.

In some surveys, respondents did not give the survey team permission to observe the handwashing facilities. These households were excluded from the denominator for calculation of the handwashing indicators.

JMP estimation rules may result in estimates that predate the establishment of a new country, area, or territory. Estimates are removed in such cases:

- No estimates are made for South Sudan before 2011
- No estimates are made for Montenegro before 2006
- No estimates are made for Timor-Leste before 2002

3.3 Ratios

In addition to the primary estimates, the JMP estimates ten ratios directly from data inputs, which refer to the proportion of people using facilities which meet some specified criteria. Ratios are not published in the JMP reports but are combined with primary estimates to produce secondary estimates which are reported.

¹² For further information see <<https://washdata.org/report/jmp-methods-task-force-report-final>>

Table 6: Ten ratios

Water	The ratio of the population using...	To the population using...
RW ₁	improved sources not exceeding 30 minutes collection time	improved sources
RW ₂	improved sources which are accessible on premises	improved sources
RW ₃	improved sources which are available when needed	improved sources
RW ₄	improved sources which are free from contamination	improved sources
Sanitation	The ratio of the population using...	To the population using...
RS ₁	improved facilities which are not shared with other households	improved facilities
RS ₂	sewer connections where wastes reach treatment plants	sewer connections
RS ₃	sewer connections where wastes reach treatment plants and are treated	sewer connections where wastes reach treatment plants
RS ₄	improved on-site sanitation facilities where wastes reach treatment plants	improved on-site sanitation facilities
RS ₅	improved on-site sanitation facilities where wastes reach treatment plants and are treated	improved on-site sanitation facilities where wastes reach treatment plants
RS ₆	improved on-site sanitation facilities where wastes are treated and disposed in of situ	improved on-site sanitation facilities

While all of these ratios are expressed as population ratios, some are constructed from volumes of wastewater or excreta: RS₃ is taken as the proportion of wastewater reaching treatment plants that is treated to at least secondary level, while RS₅ is the proportion of excreta from on-site sanitation systems reaching treatment plants that is treated.

Ratios that are used to calculate ‘safely managed’ indicators (RW₂, RW₃, RW₄, RS₂, RS₃, RS₄, RS₅ and RS₆) are calculated using JMP estimation rules, as for primary estimates. Other ratios used for ‘basic’ indicators (RW₁, RS₁) are calculated using simple averages.

Special considerations

Because separate regressions are made for each ratio, and different indicators may derive from different data sources, it is possible for ratio RW₂ to exceed RW₁. In such cases RW₁ is set as equal to RW₂ since all supplies accessible on premises are also accessible within 30 minutes.

3.4 Secondary indicators: drinking water

While nine primary indicators are calculated using JMP regression rules, the remaining 17 secondary indicators are calculated using combinations of primary indicators and ratios.

The population using **non-piped improved drinking water sources (W₃)** is calculated by subtracting piped water sources W₂ from all improved water sources W₁. Similarly, the population using **unimproved drinking water sources W₄** is calculated as the population that doesn’t use either an improved source W₁ or surface water W₅.

A **basic drinking water service (W₇)** is defined as use of an improved supply which requires no more than 30 minutes for collection (including the time required for roundtrip travel, any queuing, and filling). This is calculated by multiplying the estimate of the population using improved sources W₁ by ratio RW₁, the ratio of the population using improved sources within 30 minutes to the total population using improved sources.

A **limited drinking water service (W₆)** is calculated as the population using improved drinking water sources which are not at least a basic drinking water service. This is equal to W₁ – W₇, or W₁*(100% – RW₁).

To calculate the proportion of the population using **safely managed drinking water services (W₁₁)**, three ratios must first be calculated: the proportion of the population using improved water supplies which are accessible on premises (RW₂), have water available when needed (RW₃), and are free from contamination (RW₄). Indicators W₈, W₉, and W₁₀ are estimated by multiplying the proportion of the population using improved drinking water sources W₁ by ratios RW₂, RW₃, and RW₄, respectively. Safely managed drinking water services is taken as the minimum of these three indicators for any given year.

Table 7: Eight secondary water indicators

Water	The proportion of the population which uses...	Means of estimation
W ₃	non-piped improved drinking water sources	$W_3 = W_1 - W_2$
W ₄	unimproved drinking water sources	$W_4 = 100\% - W_1 - W_5$
W ₆	improved water sources requiring over 30 minutes (limited drinking water services)	$W_6 = W_1 * (100\% - RW_1)$
W ₇	improved water sources requiring up to 30 minutes (basic drinking water services)	$W_7 = W_1 * RW_1$
W ₈	improved water sources which are accessible on premises	$W_8 = W_1 * RW_2$
W ₉	improved water sources which are available when needed	$W_9 = W_1 * RW_3$
W ₁₀	improved water sources which are free from contamination	$W_{10} = W_1 * RW_4$
W ₁₁	safely managed drinking water services	$W_{11} = \text{minimum}(W_8, W_9, W_{10})$

Data gaps

If W₁₀ can be calculated, but either W₈ or W₉ cannot, an estimate for safely managed will be made taking the minimum of the two available indicators.

3.5 Secondary indicators: sanitation

Improved latrines or other improved on-site facilities (S₄) are calculated by subtracting sewer connections S₂ and septic tanks S₃ from all improved sanitation facilities S₁. Similarly, **unimproved sanitation facilities (S₅)** are calculated as the population that does not use either an improved facility (S₁) or practise open defecation (S₆).

If there is no information on the type of on-site sanitation facility used then these are all considered to be “improved latrines and other” (S₄) unless it is known that the majority of on-site facilities are septic tanks (e.g. in Europe).

A **basic sanitation service (S₈)** is defined as use of an improved sanitation facility which is not shared with other households. This is calculated by multiplying the estimate of the population using improved sanitation facilities (S₁) by ratio RS₁, the ratio of the population using improved facilities which are not shared to the total population using improved facilities.

A **limited sanitation service (S₇)** is calculated as the population using improved sanitation facilities which are shared. This is equal to S₁ – S₈, or S₁*(100% – RS₁).

Safely managed sanitation services (S₁₂) is calculated as the sum of three separate indicators:

1. The proportion of the population using sewer connections where wastes reach treatment plants and are treated (S₉): calculated by multiplying the population using sewer connections (S₂) by ratio RS₂ and RS₃
2. The proportion of the population using on-site sanitation facilities where wastes reach treatment plants and are treated (S₁₀): calculated by multiplying the population using on-site sanitation (S₃+S₄) by ratios RS₄ and RS₅
3. The proportion of the population using on-site sanitation facilities where wastes are disposed in situ (S₁₁) : calculated by multiplying the population using on-site sanitation (S₃+S₄) by ratio RS₆

Table 8: Eight secondary sanitation indicators

Sanitation	The proportion of the population which uses...	
S ₄	improved latrines or other on-site improved facilities	$S_4 = S_1 - S_2 - S_3$
S ₅	unimproved sanitation facilities	$S_5 = 100\% - S_1 - S_6$
S ₇	improved sanitation facilities which are shared (limited sanitation services)	$S_7 = S_1 * (1 - RS_1)$
S ₈	improved sanitation facilities which are not shared (basic sanitation services)	$S_8 = S_1 * RS_1$
S ₉	sewer connections where wastes reach treatment plants and are treated	$S_9 = S_2 * RS_2 * RS_3$
S ₁₀	on-site sanitation facilities where wastes reach treatment plants and are treated	$S_{10} = (S_3 + S_4) * RS_4 * RS_5$
S ₁₁	on-site sanitation facilities where wastes are disposed in situ	$S_{11} = (S_3 + S_4) * RS_6$
S ₁₂	safely managed sanitation services	$S_{12} = S_9 + S_{10} + S_{11}$

Data gaps

In households using sewer connections, in order for excreta to reach a treatment plant they must first be contained in piped which actually reach a sewer network, and then be transported through the sewer to the treatment plant without leaking out or being discharged. In the absence of any data, the JMP assumes that containment and transport for excreta in sewer systems are each 100%.

Similarly, on-site sanitation systems (including septic tanks and improved latrines) are assumed to fully contain the excreta flushed into them, unless there is evidence to the contrary. However, when excreta are known to be emptied from on-site sanitation systems (e.g. through a question in a household survey), but there are no data on transport to a faecal sludge treatment plant, the JMP assumes that the removed excreta *are not* transported and treated.

Data on management of wastewater in sewer systems are normally available from administrative sources such as utilities and regulators. In contrast, some data on management of on-site sanitation systems may come from households (e.g. reported emptying of tanks) while some may come from service providers (desludging companies, treatment plant operators). Frequently data are available from one but not the other of these types of sources. If data are available from the sanitation type which is used by the majority of the population, then an assumption is applied to the non-dominant sanitation type in order to make an estimate for safely managed sanitation services.

The assumptions are summarized in the following decision tree:

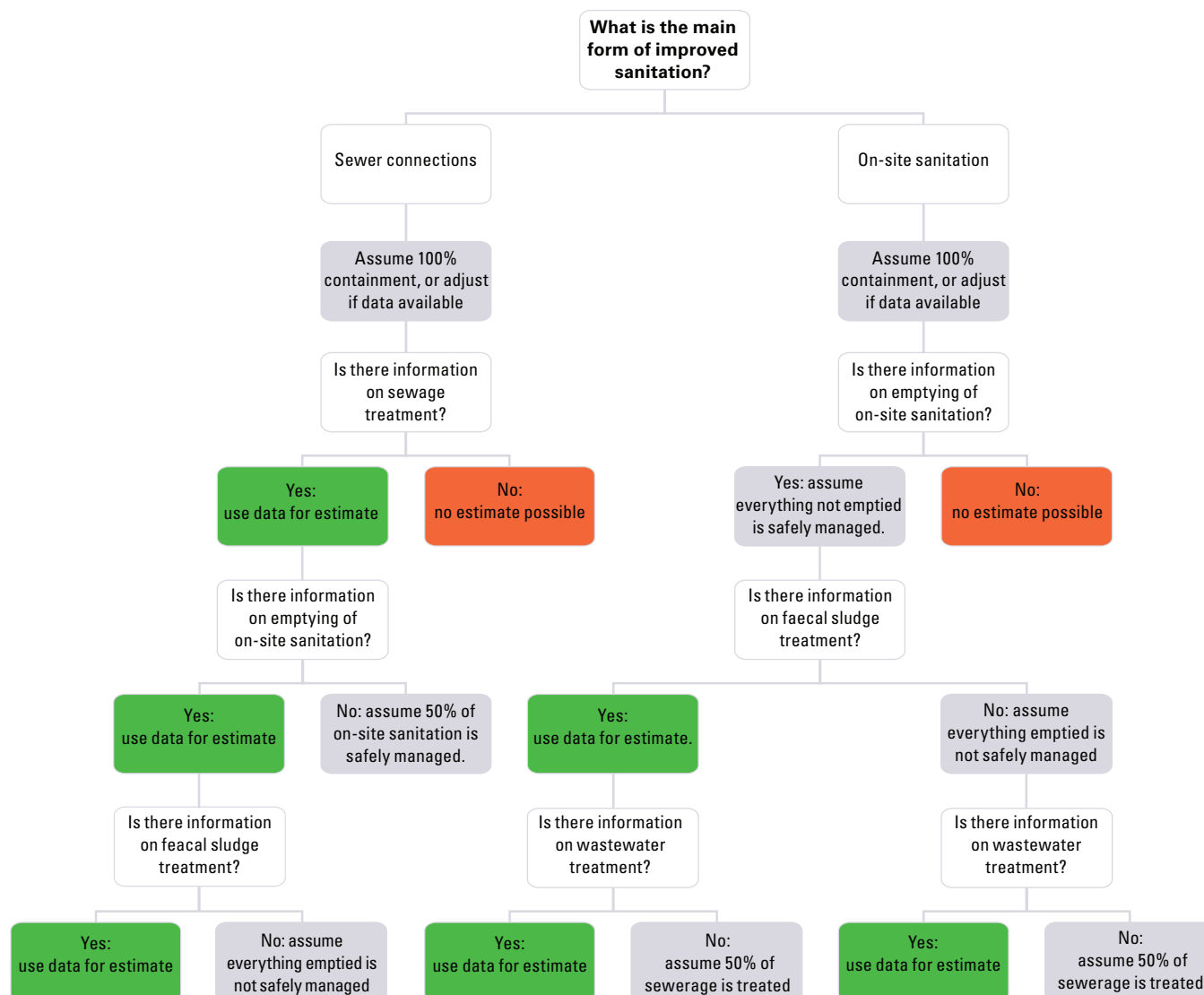


Figure 8: Decision tree for estimation of safely managed sanitation services

3.6 Secondary indicators: hygiene

The JMP uses data on the observation of handwashing facilities with water and soap, typically available in MICS and DHS surveys. Any surveys recording observation of handwashing facilities is included in the JMP database and JMP regression rules are applied to estimate the proportion of the population with a handwashing facility, as well as the proportion with a handwashing facility with water and soap.

Table 9: One secondary hygiene indicator

Hygiene	The proportion of the population which has...	
H ₃	a handwashing station in the home lacking soap and/or water (limited handwashing)	H ₃ = H ₂ - H ₁

3.7 Annual Rates of Change

In order to calculate an annual rate of change (ARC) for different indicators, the difference between two estimates

from different years is divided by the number of years separating the estimates:

$$ARC_x = \frac{P_{x,y2} - P_{x,y1}}{y2-y1}$$

Where

ARC_x is the annual rate of change for indicator x

P_{x,y1} is the estimate for indicator x in reference year y1, in percentage points

P_{x,y2} is the estimate for indicator x in reference year y2, in percentage points

The 2017 JMP Update calculated ARCs for three variables (basic drinking water services, basic sanitation services, and open defecation) for national, rural and urban populations, using 2000 and 2015 as reference years. Where estimates were lacking for either reference year, no ARC was calculated. A negative ARC represents a

declining proportion of the population reflected by the indicator. In the case of open defecation, a negative ARC represents progress towards elimination of open defecation.

3.8 Wealth quintiles

Inequalities in access to basic water, sanitation and hygiene services were assessed by household wealth quintile using a customised wealth index.

The JMP calculated customised wealth quintiles for around 80 countries for which microdata and DHS or MICS wealth quintile syntaxes were available. The wealth quintiles were recalculated by removing water, sanitation and hygiene variables based on the findings of a methodological paper¹³ which suggested that including these variables may lead to overestimation of inequalities in access. In addition to calculating quintiles of the national population, estimates were also prepared for quintiles of the urban and rural population.

For each of these countries estimates were made for rungs in the JMP ladder up to “at least basic” following the standard procedures for calculating these indicators. Due to the limited number of surveys for which wealth quintiles could be recalculated, trends were not estimated for the 2017 Update. In most cases, the results are based

on the findings of individual surveys but in some cases ratios (e.g. for protected wells) were applied from one survey to another.

3.9 Subnational

Inequalities in access to basic water, sanitation and hygiene services were also assessed by sub-national region for countries with data on the customised wealth quintiles. Sub-national regions reported in household surveys typically refer to the administrative regions below national level (known as admin-2) which are known as Divisions, Provinces, Regions etc. In some countries the surveys include a smaller number of regions (e.g. North, South, East, West) which may not correspond exactly with the administrative boundaries. The number of sub-national regions varies considerably between countries (from 3-4 to over 30) and gaps in coverage are not directly comparable.

For each country estimates were made for rungs in the JMP ladder up to “at least basic” following the standard procedures for calculating these indicators. Due to the limited number of surveys for which sub-national regions were calculated, trends were not estimated for the 2017 Update. In most cases, the results are based on the findings of individual surveys but in some cases ratios (e.g. for protected wells) were applied from one survey to another.

13 Martel et al (2016) Review of options for reporting water, sanitation and hygiene coverage by wealth quintile. <<http://uni.cf/washquintilereview>>

4.

Country consultation

In 2016, the JMP launched a “data drive” to collect data that could be used to complement data collected through its internal mechanisms (Section 1). Informal requests were sent out through UNICEF and WHO regional and country offices, WHO and UNICEF email lists, and the JMP website.

Following the data drive, preliminary estimates were produced and formally sent to countries for consultation and review. Country files, a cover memo and a guidance note were distributed to WHO Country Offices through Regional Offices, and directly to UNICEF Country Offices and Regional Offices, in February 2017. Countries were requested to focus on three questions in the consultation:

- a. Is the country file missing any relevant national sources of data on drinking water, sanitation and hygiene to get a better estimate?
- b. Are the data sources listed considered reliable and suitable for use as official national statistics?
- c. Is the JMP interpretation and classification of the data extracted from national sources accurate and appropriate?

Countries were requested to provide feedback within six weeks, by the end of March 2017. In some cases extensions were requested and made until mid-April. Multiple countries requested clarifications or corrections,

or provided additional data during the consultation period. By the end of the consultation feedback had been received from national stakeholders in over 100 countries, areas and territories.

For the purpose of global monitoring and comparison the JMP seeks to apply the same rules to all countries and for this reason some common requests from countries could not be accommodated, including:

- To use national estimates of population and urban/rural distribution. The JMP instead uses population data published by the UN Population Division.
- To assume that 100% of the population in high-income countries has basic handwashing facilities at home. In the absence of data the JMP did not make any estimates of handwashing, but will work to develop suitable proxies for future reports.
- To make provisional estimates based on partial datasets or small studies which did not meet the minimum threshold for population coverage required under the JMP rules described above.

WHO and UNICEF endeavoured to consult with all countries and respond to feedback received, and to give explanations in cases where JMP definitions or methods differed from those used or recommended by national stakeholders.

5. Calculation of regional and global estimates

Estimates of WASH services at the country level are aggregated to make regional and global estimates, following a consistent set of procedures.

5.1 Population data

The JMP uses population data from the most recent report of the United Nations Population Division, which may differ from national statistics. For the 2017 JMP Update national population figures were taken from the World Population Prospects 2015 revision¹⁴.

5.2 Imputations for countries lacking data

In order to produce estimates for regional or global levels, imputed estimates are produced for countries lacking data for one or more indicators. Imputed estimates are never shown at the national level, but are used for aggregation. Aggregated estimates are only made when the population with data coverage exceeds a specified level.

Imputations are made by calculating the regional population-weighted average for indicators, within a set of “master regions”. The master regional grouping should be selected to group countries that are as similar as possible, so that imputations are likely to reflect actual values. For the 2017 report, the UN Statistical Division’s M49 Level 2 classification¹⁵ was used as the master regional grouping. (See Table 10)

5.3 Regional and global estimates

Basic services

Regional estimates for basic drinking water, sanitation and hygiene services¹⁶ are calculated by summing up the actual or imputed estimates for each country, area or territory in the region, provided data are available for at least half (50%) of the relevant population within the region.

Table 10: M49 regional groupings

M49 Level 2 Region	Number of countries, areas and territories	2015 population (thousands)
Australia and New Zealand	2	28,498
Caribbean	26	43,199
Central America	8	172,740
Central Asia	5	67,314
Eastern Africa	20	394,477
Eastern Asia	7	1,612,287
Eastern Europe	10	292,943
Melanesia	5	9,623
Micronesia	7	526
Middle Africa	9	151,952
Northern Africa	7	223,892
Northern America	5	357,838
Northern Europe	13	102,358
Polynesia	9	684
South America	14	418,447
South-eastern Asia	11	633,490
Southern Africa	5	62,634
Southern Asia	9	1,822,974
Southern Europe	16	152,348
Western Africa	17	353,224
Western Asia	18	257,231
Western Europe	9	190,794
Total	232	7,349,472

The corresponding populations are then summed for different regional groupings used in this report. The 2017 report publishes estimates for the global region, the SDG regions, and three thematic groupings: Small Island and Developing States, Landlocked Developing Countries, and Least Developed Countries. The same methodology can be used to calculate regional estimates for other regions, such as WHO or UNICEF regions. Global estimates are made by directly aggregating country (and imputed country) estimates, not by aggregating regional estimates.

¹⁴ POP/DB/WPP/Rev.2015/F0-1, accessed June 18, 2017 from <https://esa.un.org/unpd/wpp/>

¹⁵ <https://unstats.un.org/unsd/methodology/m49/>

¹⁶ Basic level indicators include W_1 through W_7 for water, S_1 through S_8 for sanitation, and H_1 through H_3 for hygiene.

In 2017 it was possible to produce total estimates for basic water and sanitation services for all SDG regions and the world. Basic hygiene estimates were made for two SDG regions.

Safely managed services

Regional and global estimates for individual elements of safely managed services¹⁷ are calculated provided (non-imputed) data are available for at least 30% of the relevant population within the region. The relevant population for the safely managed drinking water indicators is the population using improved drinking water sources. The relevant population for the safely managed sanitation indicators is the population using the different types of improved sanitation facilities: for indicator S₉ the population with sewer connections, and for indicators S₁₀ and S₁₁ the population using improved on-site sanitation facilities.

For safely managed drinking water services, the proportions of the improved population using drinking water sources that are accessible on premises, available when needed and free from contamination are calculated as weighted averages amongst populations using improved drinking water sources. The resulting ratios are multiplied by the proportion of the population using improved drinking water sources in each region. Following the approach taken for countries, the proportion of the population using safely managed drinking water services is then calculated at regional and global levels by taking a minimum of the three elements, or of two elements if either accessibility or availability is missing. These proportions are calculated separately for urban and rural areas and, where possible, a weighted average is made of rural and urban populations to produce total estimates for the region or world.

In 2017 it was possible to produce total estimates for safely managed drinking water services in four out of eight SDG regions. Urban estimates could be made for six of the regions, while rural estimates were made for one region. Global estimates were made for total, urban and rural populations.

For safely managed sanitation services, regional estimates are only calculated if there are data on the management of the dominant form of improved sanitation (sewer connections or on-site systems) for at least 30% of the relevant population (i.e. the population using sewer connections or on-site improved sanitation systems). The population using sewer connections is used to weight estimates of the proportion of wastewater treated, while the population using improved on-site facilities is used to weight estimates of the proportion of the population with excreta disposed in situ. Where data coverage for the non-dominant form of sanitation is below 30%, estimates are based only on the dominant form of sanitation.

Regional and global estimates of the population using safely managed services are then calculated by separately summing the populations with on-site and sewer managed services. Where data coverage for the relevant population is above 30% in both rural and urban areas, a weighted average is used to produce total regional and global estimates.

In 2017 it was possible to produce total estimates for safely managed sanitation services in five out of eight SDG regions. Urban estimates could be made for four of the regions, while rural estimates were made for three regions. Global estimates were made for total, urban and rural populations.

¹⁷ Including W₈ through W₁₁ for water, and S₉ through S₁₂ for sanitation.

Annex

JMP Country Files

The Excel country files contain sheets for recording the data inputs available for water, sanitation, hygiene, population and wastewater. The first four are connected to two additional sheets – the summary sheet, used for charts, and a hidden sheet “Chart Data”.

A statistical analysis software package (Stata 14.0) imports data from the “Chart Data” sheet for all the 232 countries/ areas/territories for which data are available, and runs the estimation model. The resulting estimates are exported back to each country file in a hidden “Regression” sheet, which in turn feeds the “Estimates” tab where estimates from 2000 to 2015 are produced for the full set of indicators. The “Estimates” tab, together with the “Summary” tab, is used to create all the charts and the ladders at country level.

