Review of WASH Severity Classification

Main Report

Suggested Citation: Bedrosian N, D'Mello Guyett L, Mitchell E, and Lantagne D. 2021. *Review of WASH Severity Classification System: Main Report.* Tufts University.

Authors: Noah Bedrosian, Lauren D'Mello Guyett, Elizabeth Mitchell, and Daniele Lantagne

Photo Credit: REACH - Initiative

EXECUTIVE SUMMARY

The absence of a recognized and standardized framework to transform complex data into actionable information has become a bottleneck in decision-making processes for the Water, Sanitation, and Hygiene (WASH) sector. To remedy this, a novel WASH Severity Classification system (WSC) has been recently developed, but not externally reviewed, that classifies the severity and drivers of WASH needs based on established standards[1]. The WSC creates a severity score by combining quantitative data (including the WASH Insecurity Score (WIS)) and qualitative data (including expert judgment and consensus).

Tufts University researchers are collaborating with REACH, with funding from UNICEF, to conduct an external review of the Alpha version of WSC protocols. The findings of the review, along with the lessons learned gathered during multiple field implementations scheduled for 2021, will inform the development of the Beta version of the WSC protocol.

The overarching goal of the following work is to answer the questions:

(1) Does the WSC measure the **appropriate** social and WASH related phenomena?
(2) How **accurately** does the WSC measure the aforementioned phenomena?

To meet this goal, we completed three linked investigations: 1) literature review of existing literature (Annex A); 2) key informant interviews with persons involved in the South Sudan WSC implementation (Annex B); and, 3) statistical analysis of Somalia data to compare results from the Alpha and Alternative 1A models of the quantitative analysis to calculate the initial WIS (Annex C).

When triangulating these results across these three linked investigations, we identified five themes that emerged from the data:

Indicators. Most of the indicators used in the WSC are well supported – in the sense they are used as indicators in the literature. Some of the remaining indicators are either difficult to support (e.g. humanitarian response specific), or slightly different that existing widely used indicators. It is recommended to review the indicators used to create the WIS, considering potential opportunities for alignment with Joint Monitoring Programme (JMP) and Global WASH Cluster Indicator Bank indicators,

adjusting some indicators and ensuring that indicators used in the WIS can be consistently obtained across countries, and highlighting key indicators noted in KIIs as important / adding indicators recommended in the KIIs.

Thresholds. There was little in the existing literature to support thresholds for any WASH scores. This indicates that the WSC is innovating by developing new methods for thresholds in WASH. As with any innovation, which will need refinement over time. In particular, two refinements are recommended: 1) As no thresholds in the literature had as many components as the WSC, it is recommended to consider simplification by reducing the number of thresholds. This simplification was also recommended by KII informants and in the statistical analysis; and, 2) Consider alignment – to the extent possible and with the understanding JMP thresholds would need to be potentially expanded and adapted to the humanitarian setting – with JMP thresholds.

Analysis. KII informants expressed a desire for the WSC, and an appreciation of the holistic model merging the quantitative WIS score with analysis of the drivers and outcomes using qualitative expert judgment and consensus. In further refinements of the analysis, it is recommended to: simplify the WIS process; formalize the process of merging qualitative and quantitative data to prevent bias; and, have clear and early communication with analysts in country. Additionally, it is recommended to recomplete the statistical analysis completed in Annex C with multiple country datasets to see if results are consistent across countries, and also complete sensitivity analyses (e.g. comparing weighting, etc.) to determine if the results obtained are robust and do not vary when analysis is completed in different ways.

Forecasting. Strong desire was expressed by informants for a forecasting feature to the WSC. Informants suggested completing forecasting by simply re-completing the WSC every 3-6 months. Please note this desire by informants to complete forecasting in this manner: 1) may not be appropriate; 2) may not be possible with available data; and, 3) indicates the informants were not aware of the forecasting protocol already part of the WSC. Additionally, no forecasting models for WASH were found it the literature. It is recommended to complete research to expand, refine, and raise awareness of the forecasting feature of the WSC, which would require research and development.

Iteration. REACH has completed two versions of the WSC, Alpha, and the improved Alternative 1a model. Based on our analysis of the Somalia data, we found the Alternative 1A version of the WSC is improved over the Alpha version; showing there have been improvements over time in the model. It is

recommended to continue to iterate to improve the Beta version, as detailed herein, with additional input from a team of people, for example REACH and UNICEF staff with experience with WSC development and implementation and familiarity with JMP and global systems.

Please note that these results are further elucidated in the Annexes to this Main Report, including references, limitations, and further discussion. Readers are referred to Annex A for the literature review results, Annex B for the key informant interview results, and Annex C for the statistical analysis.

To come back to our original research questions, the answer to both is: the WSC is on the right track to be appropriate and accurate, and further iteration is needed to more fully reach these goals before rolling out the WSC internationally. This is to be expected with any development of a classification system, and planning for a WSC Beta version was already underway. We thus recommend a team of REACH and UNICEF staff work over the next 3-6 months to align, test, and validate the quantitative and qualitative portions of the WSC in REACH's development of the Beta version of the WSC before further WSC roll-out is completed.

We thank you for the collaboration, and look forward to seeing the Beta version of the WSC.

Contents

List of Figures

1 Introduction

The absence of a recognized and standardized framework to transform complex data into actionable information is a bottleneck in decision-making processes in the Water, Sanitation, and Hygiene (WASH) sector. To remedy this, a novel WASH Severity Classification system (WSC) has been recently developed that classifies the severity and drivers of WASH needs based on established standards[1]. The first version of the WSC protocols (the Alpha version) was designed in 2020, and this Alpha version was implemented in several countries in 2021. UNICEF and the GWC, in partnership with REACH, are spearheading this effort. The WSC is part of the Road Map 2020 – 2025 strategy for the humanitarian WASH sector[3].

Tufts University researchers are collaborating with REACH to conduct an external review of the Alpha version of WSC protocols. The findings of the review, along with the lessons learned gathered during multiple field implementations scheduled for 2021, will inform the development of the Beta version of the WSC. The overarching goal of the work was to answer the questions:

(1) Does the WSC measure the appropriate social and WASH related phenomena; and,

(2) How accurately does the WSC measure the aforementioned phenomena?

These two questions were divided into specific goals, and evaluated using four tools (Figure 1). In this main report, methods and results from each tool, cross-cutting themes, and conclusions and recommendations are presented.

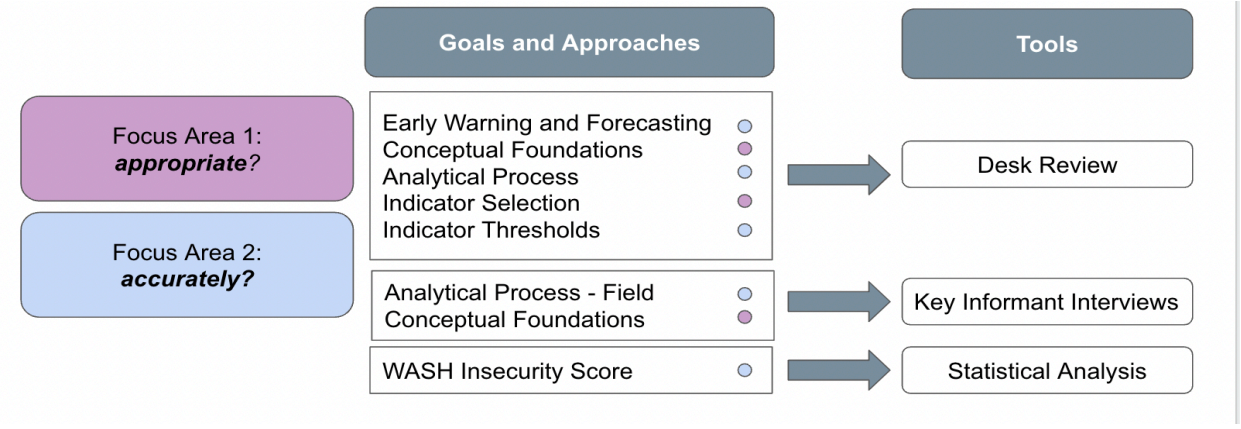


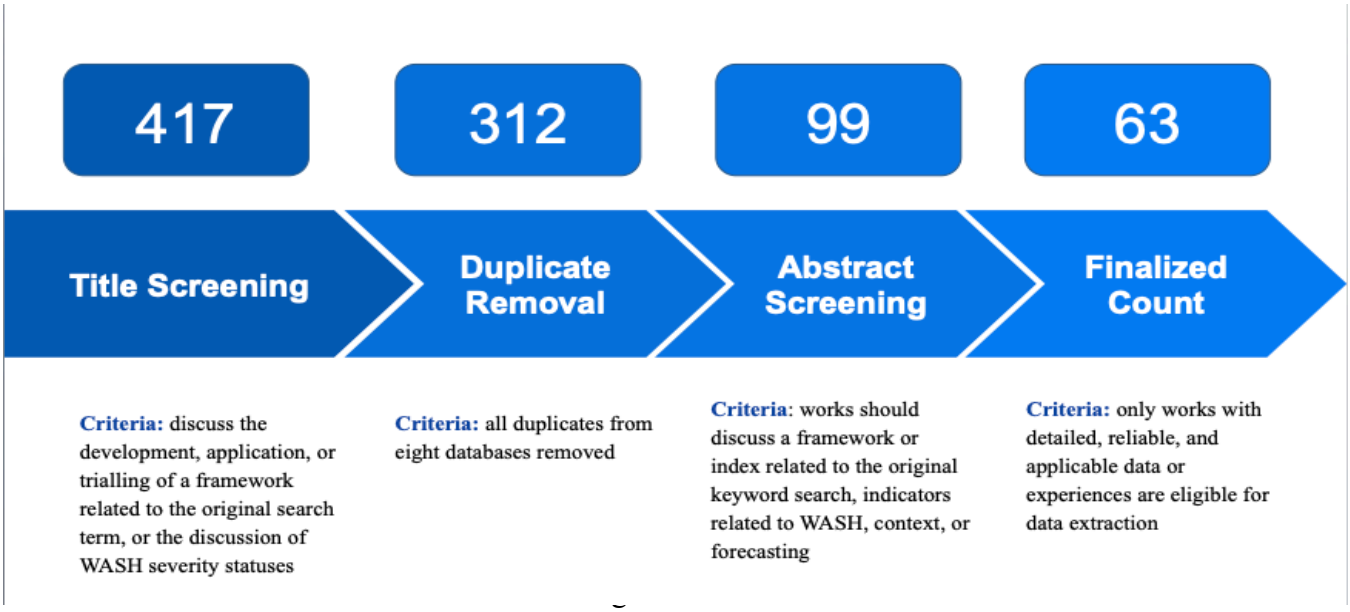
Figure 1: Mapping of Focus Areas to Evaluative Tools

2 Literature Review Methods and Results

We first conducted a literature review, including: 1) keyword and database selection; 2) title, abstract, and full text screening; and, 3) data extraction and analysis (Annex A). Each is further described below:

- In eight different databases, including both peer-reviewed manuscripts and agency documents, a list of pre-tested keywords was entered one by one, and the first 200 entries were title screened. Titles were selected if they pertained to the development, application, or trialing of a framework related to the search terms, or the discussion of WASH severity statuses. Once title screening was completed, each accepted piece of work was abstract and full-text screened.
- Following the compilation of a final list of works in Excel, each source was systematically categorized by focus area, indicator type, scale of framework, and goal type. In order to develop focused recommendations for the WSC, data were then extracted in regards to: 1) indicator selection analysis; 2) indicator threshold analysis; 3) early warning and forecasting, and, 4) emerging topics.
- Overall, 417 works were accepted through title screening, and 63 works had their data extracted. Of those 63 works, with works being able to meet multiple criteria, topics ranged from WASH (36), water security (52), sanitation (32), and hygiene (30). Indicator measurement types were continuous (44), categorical (21), or binary (8), and goals were monitoring (28), evaluation (10), identifying determinants (3), planning (2), or management (1).

Figure 2: Literature Review Screening Process and Results



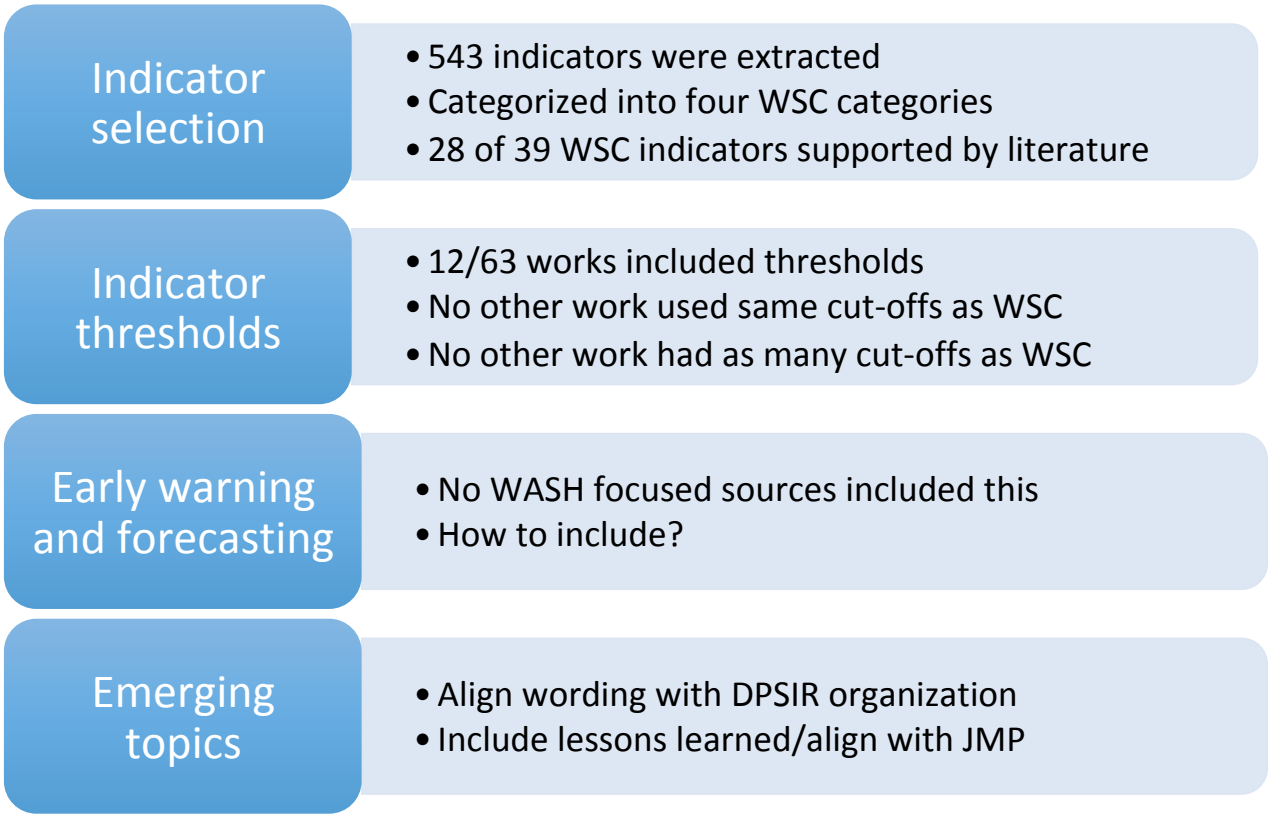
Indicator Selection Analysis. Trends in indicator inclusion were subdivided into the four categories found in the WSC: Contributing Factors, Community/Area Impact, Household Experience Outcomes, and Public Health Outcomes. In total, 543 distinct indicators were found. Based on the data available in the literature, we were able to show that 28 of 39 WSC indicators are also used in the literature. Additionally, we were able to identify indicators that are not shown to be used in the literature and indicators that can be added.

Indicator Threshold Analysis. The threshold analysis returned fewer documents than the indicator analysis, as only 12 of 63 works had indicators with individual thresholds. Additionally, many thresholds had differing numbers of cutoff levels, ranging from 2-5, which makes these data points difficult to compare. No other works found used the same severity cutoff definitions as the WSC, and thus drawing conclusions from the extracted thresholds is difficult. Thus, REACH's work in the threshold analysis is innovative and novel.

Early Warning and Forecasting. To analyze the feasibility of early warning and forecasting with the WSC, each work included in the analysis had their goals extracted. These goals included: measuring determinants, monitoring, evaluation, planning, and management. There was one work that had the goal of early warning and forecasting: the IPC Acute Food Severity Classification, of which the WSC is heavily based on this. While the IPC and WSC are very similar, there were no WASH focused sources found during the desk review that had the goal, or possibility, of early warning and forecasting. As the literature lacks recommendations for forecasting, any future REACH work to expand the existing forecasting model in the WSC would be novel, and need research and development.

Emerging Topics. Lastly, two emerging topics were identified in the review: discussion of the analytical framework and lessons to be learned from the JMP. The main recommendations were to align the wording of the indicators with Driving Forces, Pressures, State, Impact, Response (DPSIR) organization, and to include lessons learned from the JMP as the WSC is heavily based on JMP indicators and thresholds. Please see Annex A for references and further information on these emerging topics.

Figure 3: Summary Results from Literature Review

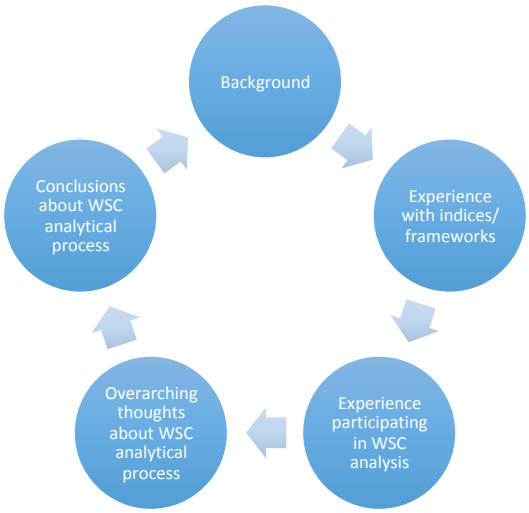


3 Key Informant Interview Methods and Results

From October to November 2021, we conducted key informant interviews on the WSC implementation in South Sudan, which was selected because it was the first country where the Alpha version of the WSC was implemented (Annex B). Questions were designed to understand informants' experiences with using the WSC in the South Sudan pilot implementation, and what the challenges, successes, or improvement areas were. Specifically, a 54-question KII guide was developed that included sections on the participant's background, their experience with other indices and frameworks, their experiences participating in the WSC analysis, their overarching thoughts on the WSC analytical process, and their conclusion about the WSC analytical process. The study protocol was approved by the Institutional Review Board at Tufts University (#STUDY00001855).

A list of people to invite to key informant interviews was provided by REACH to Tufts researchers. This list included, at REACH's request, REACH staff involved in implementing the WSC in South Sudan. Invited informants included analysts, workshop organizers, and facilitators. Interviews were transcribed, coded, and emergent themes identified. Overall, 14 informants were interviewed from the South Sudan WSC implementation, and 16 key themes emerged from the KIIs. In general, informants were satisfied with the WSC and noted key strengths, weaknesses, and areas of improvement that should be incorporated moving forward. Please note no differences in themes were identified between REACH and non-REACH staff.

Figure 4: Key Informant Interview Questions



Please note that the following text is what the informants stated. At the end of this section, we will summarize areas where informants and the existing WSC are not in alignment, indicating there are training needs, as the people implementing the WSC do not fully understand the WSC.

The informants noted that the indicator selection protocol was comprehensive and that no indicators should be removed. The key indicators identified by informants were the core WASH indicators, public health outcomes, water access and quantity indicators, nutrition indicators, and sanitation indicators. Some indicators that informants wanted to see added were including the national budget towards WASH services, looking more at menstrual hygiene management, measuring or forecasting climate change impacts, measuring the access to WASH services, a specific indicator section with thresholds on open defecation, more public health indicators, more indicators on the quality of water, coping strategies, and a measure of gender-based violence and inequities in WASH access. In terms of indicator thresholds, many informants explicitly noted that the subjective and context-based thresholds were one of the strengths of the WSC and allowed for better contextualization of the data. But some informants said that the differing interpretations of the thresholds could lead to poor or inaccurate comparability across countries if the definitions shift too much.

In terms of early warning and forecasting, informants felt that the WSC could effectively be used in this area, but more testing and trialing is needed. Specifically, the WSC could only be used for this purpose if the analysis is performed in region every six months or shorter, as to identify changes in the WASH severity of the region and to pin down the responsible drivers. But this is only possible if new data is made available during that time period. If new data takes multiple years to be produced and obtained, the early warning and forecasting system elements would not be as valuable and would have to be adjusted to accommodate the large time between analyses.

The main strengths that informants identified were the holistic nature of the WSC, and the incorporation of consensus building, local ownership of results, and the utilization of expert opinion in conjunction with other data sources. Some informants found that the in-region comparability was the major strength, while some found that the global comparability was the major strength of the WSC. Many informants also found the severity scale to be very helpful in the analysis. Most informants found that the WSC was a great tool for developmental planning and could be used to effectively coordinate WASH responses to a wide range of relevant contexts.

Informants also pointed out specific weaknesses within the WSC. While both expert opinion and consensus building were found and noted to be major strengths and key parts of the WSC, informants found that in certain settings, louder individuals could dominate the discussion and override the data or experiences of other members. Additionally, the informants stated a balance needs to be struck on the role of lived experiences versus actual data, as currently there is no clear balance between the two ideas, both of which are integral to the WSC. In terms of other weaknesses, some informants noted that newness, complexity, non-intuitive organization, dependency of data availability, and non-comparability either globally or in-region were all areas that could be improved.

Lastly, informants identified key areas of improvement and other recommendations. A few informants noted that simplification could help improve the WSC – by reducing repetitive steps, by reducing the lengths of the worksheets, or even by pulling in local analysts before the WSC starts to brief them better and allow them to be part of the data collection and pre-analysis process. The informants also stated that: 1) analysis should be performed more frequently, depending on the availability of new and accurate data for the regions; 2) specialists from other non-WASH areas, such as public health or nutrition experts, should be brought in to provide different perspectives; and, 3) vulnerable population groups should be clearly mentioned and discussed within the analysis process to ensure that the WSC correctly identifies these vulnerable groups.

Based on the key informant interview results found above, we recommend the following improvements to the WSC:

- *Modification of Expert Opinion and Consensus.* To increase clarity, equity in discussions, and reduce confusion around the role of expert opinion and consensus, a few changes could be made. During training, facilitators could be taught how to manage conversations where certain group members are dominating discussions or use anonymous forms or written feedback so that every participant gets to share their opinion. When expert opinion is used to inform decisions, remove data, or add clarification, that should be explicitly noted in outputs so other stakeholders or officials understand where that piece of data or conclusion is coming from. Additionally, there could be a limit to the amount that informants can manipulate or override data, which will have to be determined in further implementations with additional testing.

- *Determining the Feasibility of Early Warning and Forecasting.* Informants felt that the potential for early warning and forecasting was one of the greatest strengths of the WSC, but that it needed to be trialed more to ensure its feasibility. To do this, more implementations could be done in countries where the analysis has already been run, as to see how the predictions that were made played out, and how new analyses differ from the original one that was completed. Additionally, informants wanted to see the analysis done on a minimum of a six-month basis so the results could continue to be relevant and usable. However, in addition to a sound protocol and competent analysts, this analysis depends on the availability of new reliable data. Thus, this is an area in need of more trialing, and contexts with more routinely available data should be used as target areas for this testing.
- *Pre-Analysis with In-Region Analysts.* Informants felt that the WSC was too complex in the beginning or that they could have provided other data sources earlier on or given more insight into the region if they were able to be involved earlier. Because of this, there could be more educational and informative materials made available to analysts well before the implementation starts. As such, more capacity building should be done to improve the time intensity of future analyses in-region. This is so the in-region team can become more familiar with the WSC before the workshop starts. Additionally, once familiar with the WSC and its pieces, these members should be given more time and WSC resources to suggest data sources that might have been missed in the pre-analysis.
- *Analysis Worksheet Simplification.* In addition to performing more pre-analysis with in-region analysts to reduce complexity, the analysis worksheet should be simplified. This does not mean that indicators should be removed, but rather that more training should be done to familiarize the participants with the worksheet and break the larger sections into more easily understandable parts. The researchers recognize that the worksheet has been simplified since the South Sudan implementation.
- *Addition of New Indicators.* During the analysis, the following areas of additional indicator inclusion were suggested by informants (please note these are reported from informants, and if informants reported them as new indicators needed and they are already in the WSC, it means either the informants didn't know they were there or they need expansion) and should be considered for addition to the WSC:
    - The national budget towards WASH services; Menstrual hygiene management;

    - The effects of climate change; Access to WASH services; Open defecation;
    - More public health indicators;
    - More indicators on the quality of water, such as biological testing parameters, water quality at the source, water quality in the home, or water quality in municipal systems;
    - Gender-based violence and gender-based inequities in WASH.
- *Defining the Goals of the WSC Outputs.* Informants found that the WSC was walking the line between two important areas –globally comparability, or in-country comparability. Due to the subjective and possibility of differing interpretations of thresholds, some informants worried that the severity scores would not be comparable globally. However, informants also felt that the contextualization of data sources through the use of expert opinion made the WSC a novel and valuable tool for in-region comparisons, prioritization, and decision-making. It would be more globally comparable if the country wide severity level had less subjective thresholds or relied on less expert opinion, while more comparable if the in-region severity levels considered the context of the regions they were in. This is an area that will require more testing, reflection, and iteration as the WSC continues to be implemented.

Please note in the above summary of informant statements there are some inconsistencies with the existing WSC protocols, for example in forecasting potential, in how the analysis is completed. In particular, the WSC development team requested a clarification to this portion of the report saying: *"the thresholds for the quantitative indicators are not designed to be subjective and should stay the same per context. Rather contextualization is expected through expert judgment. This misunderstanding of the KIs is further demonstrative of the need for training on the different WSC protocols, as elaborated below".* The fact informants did not accurately understand the existing WSC indicates a need for additional training of those actually implementing the WSC so they understand the WSC. Additionally, please note these interviews were conducted in only one implementations setting, and it is recommended to recomplete these interviews in other country settings. Lastly, further detail is available in Annex B.

4 Statistical Analysis Methods and Results

Across the last year, several countries have adopted and piloted the WSC methodology to assess national and local severity of WASH needs. Considering the pilots and future use, it is necessary to analyze how certain parameters of the WSC methodology are validated and influence results. The WSC methodology for calculating severity can be broken down into two overarching steps:

1. Calculation of the WASH Insecurity Score (WIS), a composite index that generates a household-level score based on a set of core WASH indicators. This provides a preliminary indication of the number of households in each WSC severity phase.
2. Comparison of the WIS in conjunction with data (both quantitative and qualitative) on other elements of the WSC Analytical Framework, using expert judgment to interrogate, explain, and adjust the score and to develop an overall phase classification for the area.

The WIS therefore serves as a quantitative foundation upon which data pertaining to other areas of the WSC Analytical Framework are then incorporated (often in a more qualitative fashion), and therefore needs to be based on quantifiable, standardized, objective indicators.

Focusing on only the quantitative WASH Insecurity Score (WIS), which underlies the overall WSC, the objectives of this study were to:

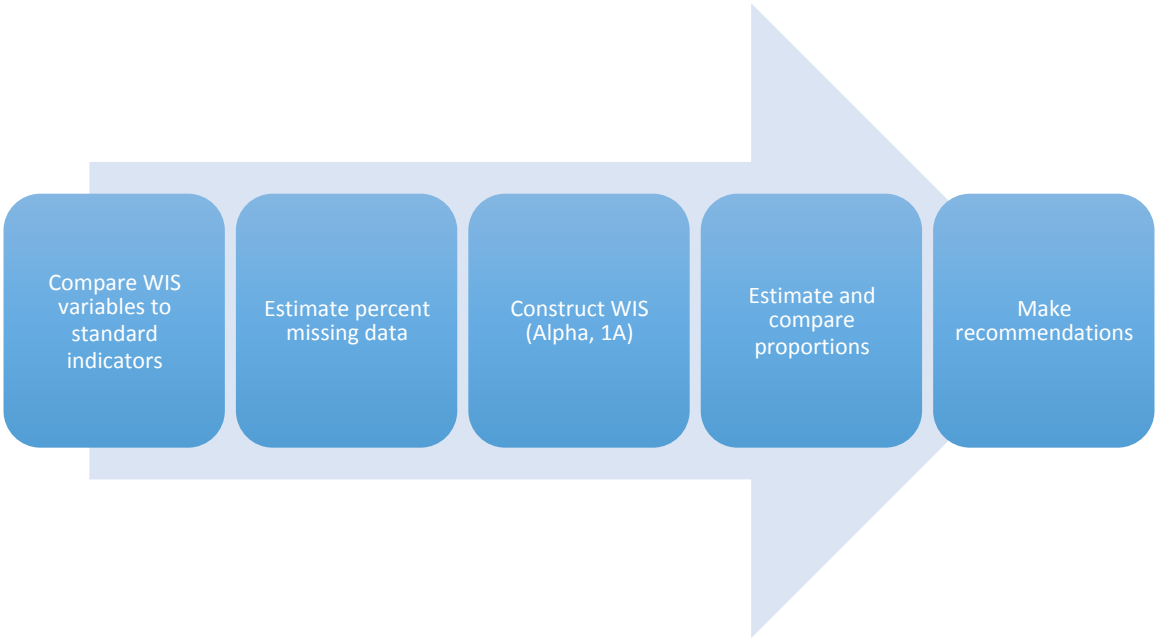
- Analyze the how the WIS has been constructed; and,
- How the model varies by inclusion of different indices.

Using the latest available survey data from Somalia, our objectives and methods were to (Annex C):

- Compare variables used by the WIS and compare their definitions to standard WASH indicators;
- Estimate the proportion of missing data;
- Construct WASH indices according to the two currently available WIS models – the original more complex initial Alpha model and a simplified Alternative 1a model that removed, based on lessons learned, the ‘water sufficiency’ indicator;

- Estimate the proportion of WASH coverage (the WIS) for one country based on Alpha and Alternative 1a including an estimate of the population with combined WASH coverage by phase quintile; and,
- Make recommendations on how to improve the model.

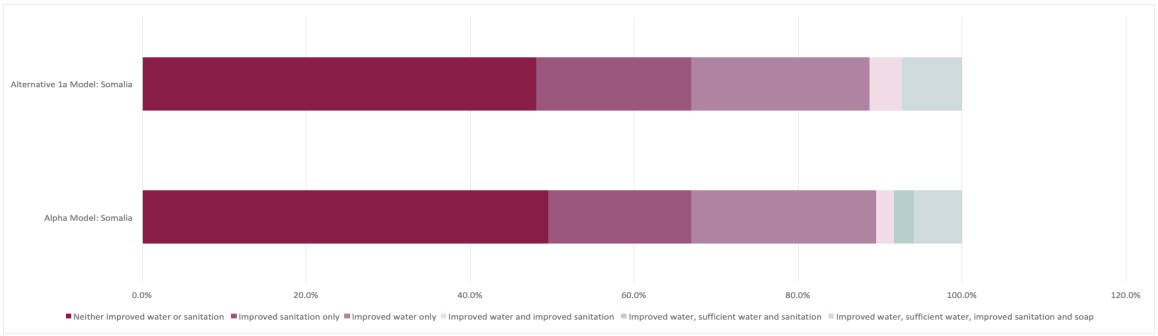
Figure 5: Methodology for Somalia Data Statistical Analysis



There were three main results from this statistical analysis in Somalia, each described below:

- There was strong correlation between the Alpha and Alternative 1a models; as such we recommend moving to the Alternative 1a model, as it is simpler.

Figure 6: Results for Somalia Data Comparison of Alpha and Alternative 1a Models



- We recommend further refining the Alternative 1a model, as planned, into the Beta model:
    - Expanding the handwashing indicator beyond soap access only;
    - Further aligning the water and sanitation indicators with JMP definitions; and,
    - Further aligning all JMCNA indicators with JMP and the Global WASH Cluster 'Indicator Bank' to ensure consistency of the indicators used across countries. Please note "alignment" in this case does *not* mean "use JMP indicators exactly". It means ensuring that indicators used in the WIS are available, comparable to other indicators, and perhaps expanded (where appropriate) to include indicators particularly relevant to the humanitarian setting.

Overall, we found strong correlation in how the two WASH Insecurity Score (WIS) models rank combined access to WASH. We found only a slight difference between the original Alpha and Alternative 1a models. With the caveat that this analysis was conducted in only Somalia and with data from only 2021, this analysis supports the decision to move away from the more complex Alpha model to the simplified Alternative 1a.

However, the Alternative 1a model does not measure the presence of a handwashing facility as a measure for hygiene. To make the WIS more in line with standardized definitions used across the WASH sector, the Alternative 1a model can be updated to include the measurement of handwashing facility alongside the presence of soap, by including:

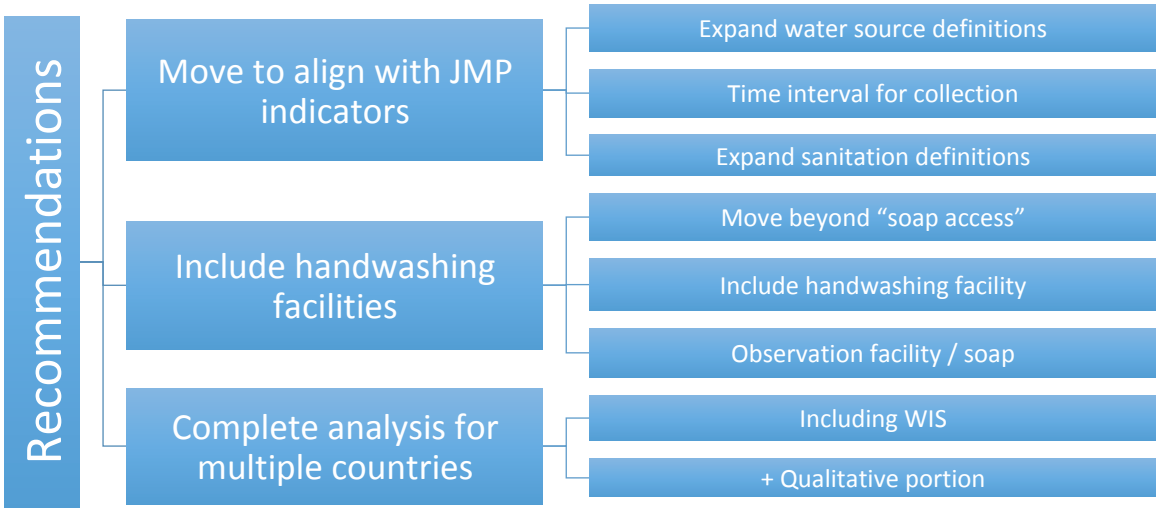
- Observation of a handwashing facility observed (fixed facility or mobile object); and,
- Observation of soap or detergent (bar, liquid, powder, paste).

Going forward, we would recommend that the list of indicators included within the WIS models, and the JMCNA/MSNA surveys that are collected to calculate WASH coverage across all countries, are updated to be consistent with JMP terminology, especially definitions of water and sanitation infrastructure. It is understood that the JMCNA indicators were taken from the Global WASH Cluster indicator bank. This recommendation therefore similarly applies to the GWC indicator bank. We recommend that the updates would include the following changes to all JMCNA/MSNA surveys, to enable consistency across countries, and results.

- Expanding the definitions for water source types to include technologies such as public taps or standpipes, springs, and packaged water;
- Reducing the intervals for time to fetch water to either "on premises or less than 30 minutes" and "more than 30 minutes", and potentially a longer timeframe to potentially expand the JMP indicator to humanitarian settings;
- Expanding the definitions for sanitation facilities to include on site sanitation such as composting toilets or septic tanks; and,
- Including a question on the urban or rural residence of the surveyed population to understand sub-national inequalities.

Lastly, we recommend this statistical analysis be completed in the future across many country data sets to ensure consistency of results.

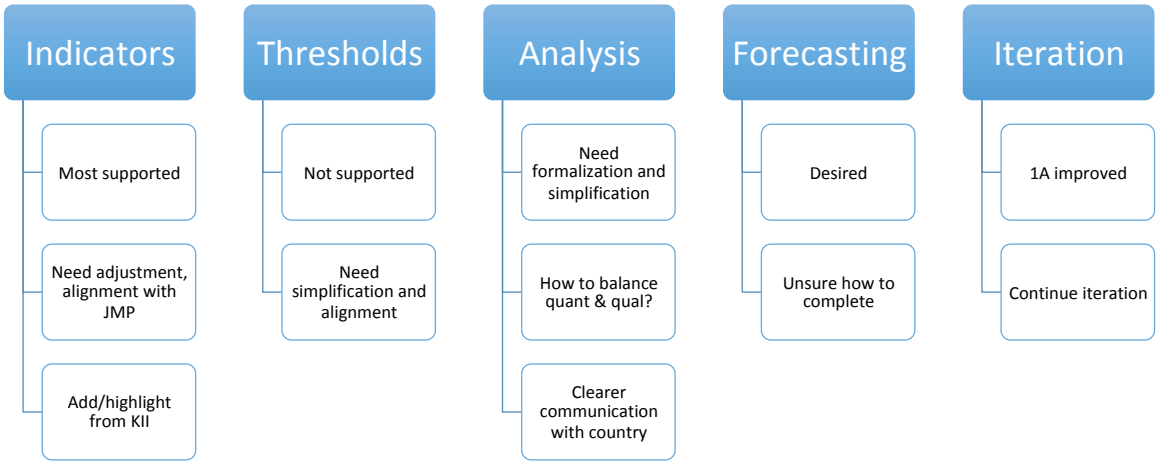
Figure 7: Recommendations from Statistical Analysis



5 Emergent Themes and Conclusions

Overall, five cross-cutting emergent themes evolved from the three investigations: indicators, thresholds, analysis, forecasting, and iteration.

Figure 8: Cross-cutting Emergent Themes from Triangulated Results



Indicators. Most of the indicators used in the WSC are well supported in the literature. Some of the remaining indicators are either difficult to support (e.g. humanitarian response specific), or slightly different that indicators used in the literature. It is recommended to:

- Review the indicators used to create the WIS;
- Consider potential opportunities for alignment with Joint Monitoring Programme (JMP) and Global WASH Cluster Indicator Bank indicators;
- Consider expanding JMP indicators to be more relevant to humanitarian contexts (for example expansion of lower segments of the ladder to include potentially some indicators from Sphere or other humanitarian response indicators);
- Determine if key indicators noted in KIIs as important or potentially to add should be added; and,
- Ensure all selected indicators can be consistently obtained across countries.

After a key set of indicators that can be obtained reliably across countries is determined, those indicators should be tested in a new statistical analysis.

Thresholds. There was little in the existing literature to support thresholds for WASH scores. This indicates that the WSC is developing new methods, which will need refinement over time. In particular, it is recommended to consider simplification of the thresholds (as recommended by KII informants and in the statistical analysis) and/or potential alignment with JMP thresholds where possible. These new thresholds should be tested in a new statistical analysis.

Analysis. KII informants expressed a desire for the WSC, and an appreciation of the holistic model merging the quantitative WIS score with qualitative expert judgment and consensus. In further refinements of the analysis, it is recommended to: simplify the WIS and WSC processes; formalize the process of merging qualitative and quantitative data to prevent bias; and, have clear and early communication with analysts in country. Additionally, it is recommended to repeat the statistical analysis completed in Annex C with the updated indicators and thresholds with multiple country datasets to see if results are consistent across countries, and also complete sensitivity analyses (e.g. comparing weighting, etc.) to determine if the results are robust. Please note some of these analyses were planned for this work, but data to complete them was not available during the time of the work.

Forecasting. Strong desire was expressed by informants for a forecasting feature to the WSC. Informants suggested simply re-completing the WSC every 3-6 months, which may not be appropriate and may not be supported by available data. Additionally, no forecasting models for WASH were found it the literature. It is recommended to complete research on developing a method and/or protocol to potentially further develop the Alpha protocol for forecasting.

Iteration. The Alternative 1A version of the WIS is improved (based on analysis of Somalia data) over the Alpha version, showing improvements over time in the model. We thus recommend a team of REACH and UNICEF staff work over the next 3-6 months to align, test, and validate the quantitative and qualitative portions of the WSC in REACH's development of the Beta version of the WSC before further WSC roll-out is completed.

To come back to our original research questions, the answer to both questions

1. Does the WSC measure the **appropriate** social and WASH related phenomena?
2. How **accurately** does the WSC measure the aforementioned phenomena?

is: the WSC is on the right track to be appropriate and accurate, and further iteration is needed to more fully reach these goals before rolling out the WSC internationally. This is to be expected with any development of a classification system, and planning for a WSC Beta version was already underway. We thus recommend a team of REACH and UNICEF staff work over the next 3-6 months to align, test, and validate the quantitative and qualitative portions of the WSC in REACH's development of the Beta version of the WSC before further WSC roll-out is completed.

We thank you for the collaboration, and look forward to seeing the Beta version of the WSC.

Acknowledgements

We would like to acknowledge the WSC team at REACH, the donors at UNICEF, the key informant interview participants, and the team of undergraduate researchers at Tufts University who transcribed the interview reports.