

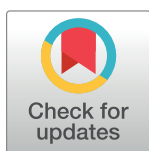
OPINION

Challenges and opportunities for advancing data-driven WASH programming: Reflections from the UNC Chapel Hill *Water and Health Conference* side event "DATA: A key for unlocking quality in WASH programming"

James Brown^{1*}, Syed Imran Ali¹, Sydney Hubbard², Laura MacDonald³, Rachel Peletz⁴, Katy Sill⁵, Mian Zhong⁶

1 Dahdaleh Institute for Global Health Research, York University, Toronto, Canada, **2** mWater, Los Angeles, California, United States of America, **3** Mortenson Center in Global Engineering & Resilience, University of Colorado Boulder, Boulder, Colorado, United States of America, **4** The Aquaya Institute, San Anselmo, California, United States of America, **5** Global Water Challenge, Arlington, Virginia, United States of America, **6** Global Health Engineering, ETH Zürich, Zürich, Switzerland

* brownje@yorku.ca



OPEN ACCESS

Citation: Brown J, Ali SI, Hubbard S, MacDonald L, Peletz R, Sill K, et al. (2024) Challenges and opportunities for advancing data-driven WASH programming: Reflections from the UNC Chapel Hill *Water and Health Conference* side event "DATA: A key for unlocking quality in WASH programming" PLOS Water 3(11): e0000316. <https://doi.org/10.1371/journal.pwat.0000316>

Editor: Guillaume Wright, PLOS: Public Library of Science, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Published: November 21, 2024

Copyright: © 2024 Brown et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The Dahdaleh Institute for Global Health Research received funding from ELRHA's Humanitarian Innovation Fund (<https://www.elrha.org/funding-support/hif/>) to support the convening of the UNC Side Event on which this article is based. The funders had no role in the design of the event or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Achieving universal and equitable WASH services requires accurate, data-driven understanding of context, needs, and evidence. Investing in data systems enables stakeholders to assess needs, identify priorities, and allocate resources efficiently. As researchers and practitioners working in low- and middle-income countries (LMICs), we have never had greater access to data. However, there is still much that we can learn about how to translate this data into action in the pursuit of more effective, equitable, and accountable WASH programming.

During the 2023 *Water & Health Conference* at the University of North Carolina, Chapel Hill, we convened a meeting of WASH researchers, practitioners, and data specialists to discuss the current state and trajectory of data-driven WASH programming in LMIC contexts. Our goals were to identify opportunities for the application of data to drive improvements in WASH programme quality, and to foster collaboration among organisations working in this space. Here we summarise four emergent themes, documenting examples of, and recommendations for, action. We aim this article at academic researchers, practitioners, and governments, particularly those involved in the design, implementation, and evaluation of WASH programmes in LMICs and in humanitarian contexts. We acknowledge that our perspectives are primarily rooted in the USA, Canada, and Europe, and recognize that a more globally inclusive dialogue around WASH data is needed to build a more comprehensive understanding of these issues.

I. We need to measure what matters

Enhancing the usefulness and relevance of WASH data begins with the careful selection of indicators that provide meaningful insight into specific programmatic questions. This requires recognising when proxy indicators fall short and instead opting for direct measures, even if harder to collect.

For example, in water supply programmes, there is mounting evidence that recontamination of treated water in intermittent water systems commonly occurs during the ‘last mile’ of collection, storage and consumption [1]. Despite this, water quality sampling is often only conducted upstream, at tapstands or the point-of-treatment, thus risking the mischaracterization of water safety at the household, where the risks of waterborne disease manifest.

To address this challenge, researchers at the Dahdaleh Institute for Global Health Research created the *Safe Water Optimization Tool* (www.safeh2o.app), a water quality modelling platform that generates site-specific and evidence-based water treatment guidance to help prevent recontamination over the critical last mile. Field experience with the tool has driven advocacy in the humanitarian WASH sector to monitor and address public health risks in a more direct way, at the household level where exposures happen.

Progress has been made developing established indicator frameworks for specific applications, such as those set by the Water Point Data Exchange (WPDx) for water infrastructure and the Joint Monitoring Programme’s standardised approach to household surveys. Moving forward, there is scope for demonstrating how these approaches could be further integrated to meet the needs of different users and contexts.

II. We need to promote accessibility of WASH data

Primary data collection is expensive and time consuming for researchers and practitioners alike, and it is often burdensome for those from whom data is being collected. Against a backdrop of scarce funding, where data collection investments are made, it is crucial to ensure that the resulting datasets are accessible—easy to find, understand, and use. Where best practices are not followed, incredible volumes of potentially useful data can end up ‘siloes’—rendered inaccessible to common data extraction techniques.

Funding requirements and institutional priorities play a role in shaping these practices. Data owners are often motivated by specific project reporting requirements rather than maximising the intrinsic value of data for broader impact. For example, a focus on quantifiable metrics and project-cycle timelines can lead organizations to prioritize capturing data to demonstrate immediate results rather than contributing to the longer-term, sustainable advancement of safe WASH access.

While meeting participants expressed frustration over siloes data, they were equally concerned about privacy and the risks of improper data sharing. Navigating the balance between preserving detailed contextual information and adhering to ethical standards presents a complex challenge, often leading to overly cautious data protection measures. There is an opportunity for the development of better guidance outlining what can be responsibly shared and how, within the framework of global data standards.

Openwashdata (openwashdata.org) is a global community that promotes data sharing in the WASH sector by advocating for and supporting data owners to prepare and publish their data. The initiative aims to build motivation for sharing through the registration of datasets as citable works, offering recognition of data collectors’ contributions.

Open data exchanges, like WPDx and the Humanitarian Data eXchange (HDX) and platforms such as mWater have a big part to play in developing data standards, acting as data repositories and providing tools to support decision-making. Of course, in many cases, national agencies or humanitarian coordination platforms have established standards, tools, and data repositories in place. Building upon and supporting these existing initiatives is imperative.

For example, the Millennium Water Alliance uploaded over 20,000 water point records from government and NGO partners in Ethiopia to WPDx [2]. Decision-support tools were

then used by local and regional government offices to prioritise interventions, justifying rehabilitation over new construction as a more efficient and impactful approach [3].

III. We can find value from integrating datasets in new ways

Detailed, field-level data that provide insight into local conditions are often needed to build understanding of needs, but these tend to be fragmented, both spatially and temporally. Coverage typically prioritizes areas that may be relatively well-served or easy to access. Achieving an unbiased assessment of needs, risks, and/or service levels requires comparable data spanning much wider areas. Although there is huge potential in the application of remote sensing, taken alone, it is unable to provide the detail required to address WASH-related questions.

Integrating datasets across regions and timeframes can fill these gaps. Comparing large-scale remote sensing data with detailed field data helps generate actionable insights. For example, the Drought Resilience Impact Program of the Mortenson Center in Global Engineering & Resilience integrates data from in-situ borehole sensors with remote sensing data to forecast local drought emergency risk over a wide area [4]. While wide area forecasting cannot replace the need for field-level assessments and monitoring, the ability to quickly identify areas where risks or needs are acute has potential to address disparities in service provision, intervention targeting, and mitigation planning. As the availability of remotely collected and derived data continues to grow, the WASH sector should embrace opportunities to situate detailed local perspectives within the macro-level environment. To ensure that these insights lead to real-world outcomes, data must be aligned with the needs of decision-makers.

IV. We need to ensure that data is used to drive action

In addition to ensuring that WASH data is relevant, accessible, and integrated, it is equally critical that data aligns with the requirements of decision-makers, particularly national and local governments that are responsible for safe WASH provision.

We must acknowledge the mismatch between researcher, practitioner and local government priorities. Engaging from the outset with government agencies, service providers, and consumers themselves is key to understanding and aligning with their priorities and processes. Developing processes for stakeholder engagement is also critical to build consensus and trust and to help guide the development and adaptation of data systems to best fit stakeholder needs. While government decision-makers may not engage directly with datasets collected or compiled by non-government actors in the short term, interim strategies like collaborative data-sharing, capacity-strengthening programmes, and phased ownership can help governments take a central role in managing these platforms. Strengthening government capacity and planning to align and integrate data systems gradually will support long-term sustainability.

Timing is also crucial: insights must be available to stakeholders in time to act. Lengthy ethics review and publication timelines can obstruct the application of research data to operational decision making; streamlining this process through prior planning and providing interim actionable analysis, where feasible, helps ensure it meets the needs of both researchers and practitioners.

As an example of how improving monitoring of water quality and effectively communicating results can drive improved water treatment by building demand for safe water, the Aquaya Institute (Aquaya) is working with small, piped water systems in Ghana and local government [5]. In this situation, the regular in-person meetings with water system operators and radio dissemination to reach consumers helped ensure that water quality data was used to inform

water treatment decisions. Currently, Aquaya is documenting these types of impacts as part of a larger, randomized controlled trial of a water quality monitoring intervention [6].

V. Final thoughts

As the WASH sector's challenges continue to be reshaped by political, technological, and environmental forces, leveraging data for effective, adaptive, and equitable programming becomes even more crucial. It's essential to cultivate a culture that sees data not as an end product but as a raw material with significant opportunity cost. This approach demands deliberate system design to support data collection, sharing, integration, and analysis if it is to meet the diverse needs of stakeholders.

Incentives and policies are key to fostering these desired shifts, including recognition of organizations that contribute substantially to shared data repositories, funding preferences for projects that demonstrate strong engagement with key decision-makers, and support to help organizations standardize, share, and distil their data into key insights. While there is an important role here for funders and policymakers, embedding a culture of data sharing and use requires ongoing sector-wide engagement, capacity building, and advocacy.

While technological advances call for greater investment in data capabilities, we must recognise the resources already being deployed to this end and the opportunities that already exist through greater collaboration and sharing.

This event highlighted many examples of new ways data is being applied to strengthen WASH programming, as well as similar frustrations faced by those advocating for more effective use of valuable data resources. Encouraged by these shared experiences, we remain deeply optimistic about data's role in WASH programme quality and accountability and keenly anticipate further shared progress and deeper collaboration in this crucial field.

Acknowledgments

We greatly appreciate Denis Macharia Muthike (Mortenson Center in Global Engineering & Resilience) for his support in conceptualizing the event; and Jason Lopez (Millennium Water Alliance) for enriching the discussion with his knowledge and experience. We are also grateful to Brian Banks (USAID) and Mohammed Bah (Sierra Leone Ministry of Water Resources) for reviewing an early draft of this article and offering their valuable perspectives. While the views expressed in this article are solely those of the authors, their contributions significantly enhanced the work.

Author Contributions

Writing – original draft: James Brown.

Writing – review & editing: James Brown, Syed Imran Ali, Sydney Hubbard, Laura MacDonald, Rachel Peletz, Katy Sill, Mian Zhong.

References

1. Ali SI, Ali SS, Fesselet JF. Evidence-based chlorination targets for household water safety in humanitarian settings: Recommendations from a multi-site study in refugee camps in South Sudan, Jordan, and Rwanda. *Water Res.* 2021 Feb; 189:116642. <https://doi.org/10.1016/j.watres.2020.116642> PMID: 33246215
2. Additional Insights from Implementing WPdx in Ethiopia [Internet]. WPDx. 2023 [cited 2024 Feb 27]. Available from: <https://www.waterpointdata.org/2023/08/28/additional-insights-from-implementing-wpdx-in-ethiopia/>
3. Selamawit Tiruneh, Mussie Tezazu. Validating the Impact: A Field Visit Assessment of WPdx Rural Decision Support Tools in Ethiopia [Internet]. WPDx. 2023 [cited 2024 Feb 27]. Available from: <https://www.waterpointdata.org/2023/07/12/validating-the-impact-a-field-visit-assessment/>

4. Fankhauser K, Macharia D, Coyle J, Kathuni S, McNally A, Slinski K, et al. Estimating groundwater use and demand in arid Kenya through assimilation of satellite data and in-situ sensors with machine learning toward drought early action. *Sci Total Environ.* 2022 Jul; 831:154453. <https://doi.org/10.1016/j.scitotenv.2022.154453> PMID: 35346702
5. Kumpel E, MacLeod C, Stuart K, Cock-Esteb A, Khush R, Peletz R. From data to decisions: understanding information flows within regulatory water quality monitoring programs. *Npj Clean Water.* 2020 Aug 24; 3(1):38.
6. REAL-Water. Evaluating Water Quality Assurance Funds in Ghana: Baseline Assessment. United States Agency for International Development (USAID); 2023. (Rural Evidence and Learning for Water Project).