SUPPLEMENTARY TOOL

Information and examples on data merge

WASH and NTD data merge implementation will be different in each country given the unique context of national systems, the programmatic and policy needs, and the timing and forum for determining budget allocation. The following can be used as a generalized checklist to merge WASH and NTD data to inform programmatic and policy decisions.

First, determine what systems exist for both WASH data and NTD data, what data elements are in each system, if the data are continuous or categorical, at what level these data are stored, and how these data can be accessed. Often, you will find data are stored at a district or subdistrict level and use a system for identification of the administrative unit (name, code system, etc.). WASH and NTD data will be merged by a common administrative unit. In order to achieve this, working with the relevant ministries or authorities (Water and Health), the level of data capture and identification names may need to be aligned.

Once the data sources and the data in each source have been identified, both data elements need to be transformed into categorical values and merged. The merged data are then ready for cross-tabulation analysis, where the relationship between NTD prevalence data and WASH data elements can be explored further. It is important to consider the policy decisions and biological plausibility between each NTD and each WASH indicator before performing this analysis.

Table 1. **WASH and NTD data merge cross-tabulation table example**

	WASH resource access **low**	WASH resource access **high**
NTD prevalence **high**	**Low/High**	High/High
NTD prevalence **low**	Low/Low	High/Low

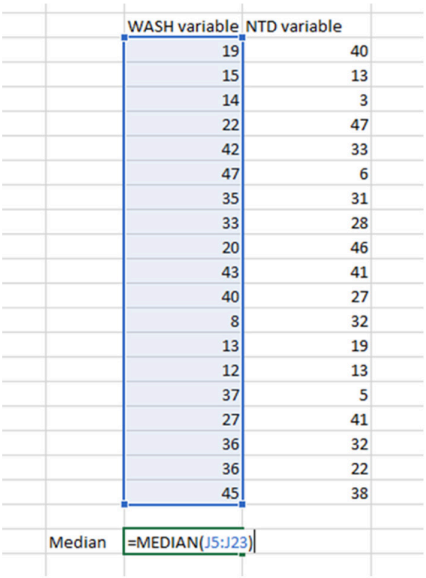
Prerequisites

To get started you will need to have both WASH and NTD data that share a common data element, such as district or implementation unit (IU) name. You will be combining these data by this common data element and recoding the data to either high or low based on some predetermined threshold or calculated value.

Preparing your data

Step 1: Create a value to determine high and low

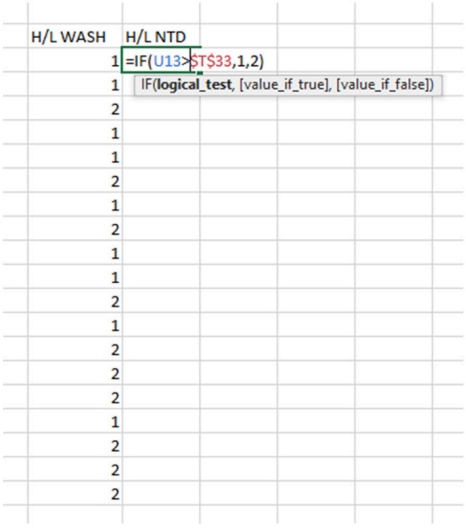
To determine if an individual value should be categorized as either high or low, you can either use a predetermined value or some calculated value, such as the median of all values. To calculate the median, you can use the median function in Microsoft Excel.



	WASH variable	NTD variable
	19	40
	15	13
	14	3
	22	47
	42	33
	47	6
	35	31
	33	28
	20	46
	43	41
	40	27
	8	32
	13	19
	12	13
	37	5
	27	41
	36	32
	36	22
	45	38
Median	=MEDIAN(J5:J23)	

Step 2: Categorize value to high or low

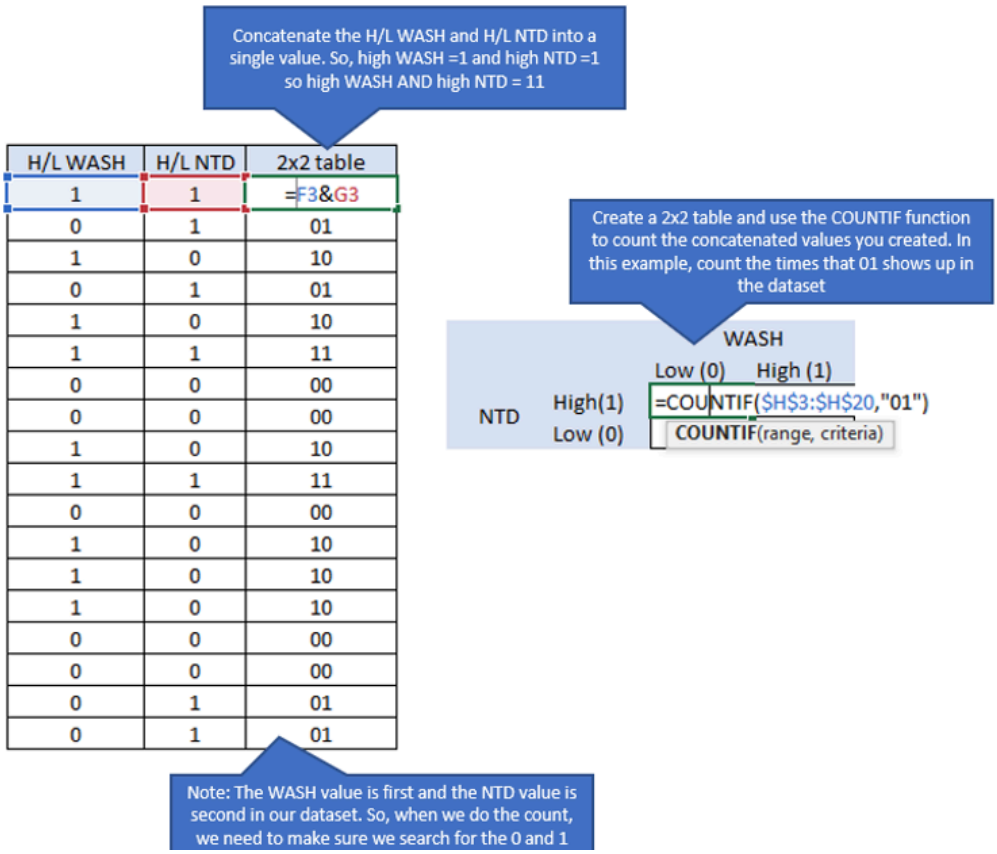
Once you have established a cut off, either by using a predetermined value or the median, you can use the IF function in Microsoft Excel to assign each value to either high or low. Below, you can see values were assigned as 1 if above the median and 2 if less than or equal to the median.



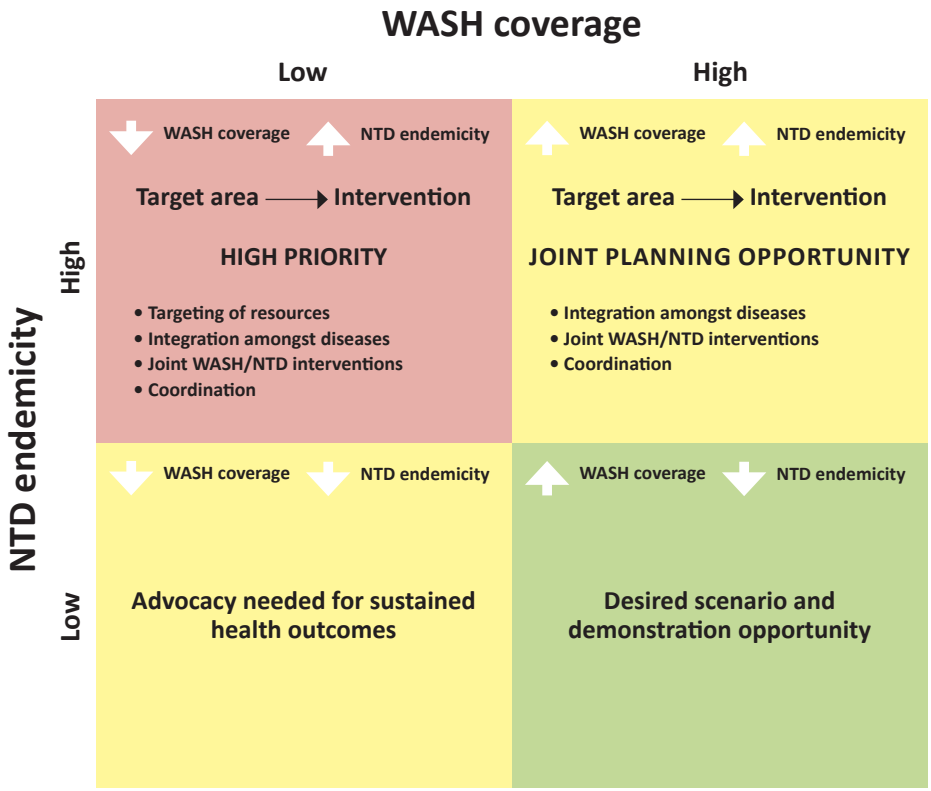
H/L WASH	H/L NTD
1	=IF(U13>ST$33,1,2)
1	IF(**logical_test**, [value_if_true], [value_if_false])
2	
1	
1	
2	
1	
2	
1	
1	
2	
1	
2	
2	
2	
1	
2	
2	
2	

Step 3: Create a cross tabulation

Now that you have assigned WASH data to high or low, and disease prevalence data have their own categories, you can create a table to assign IUs based on their combined status.



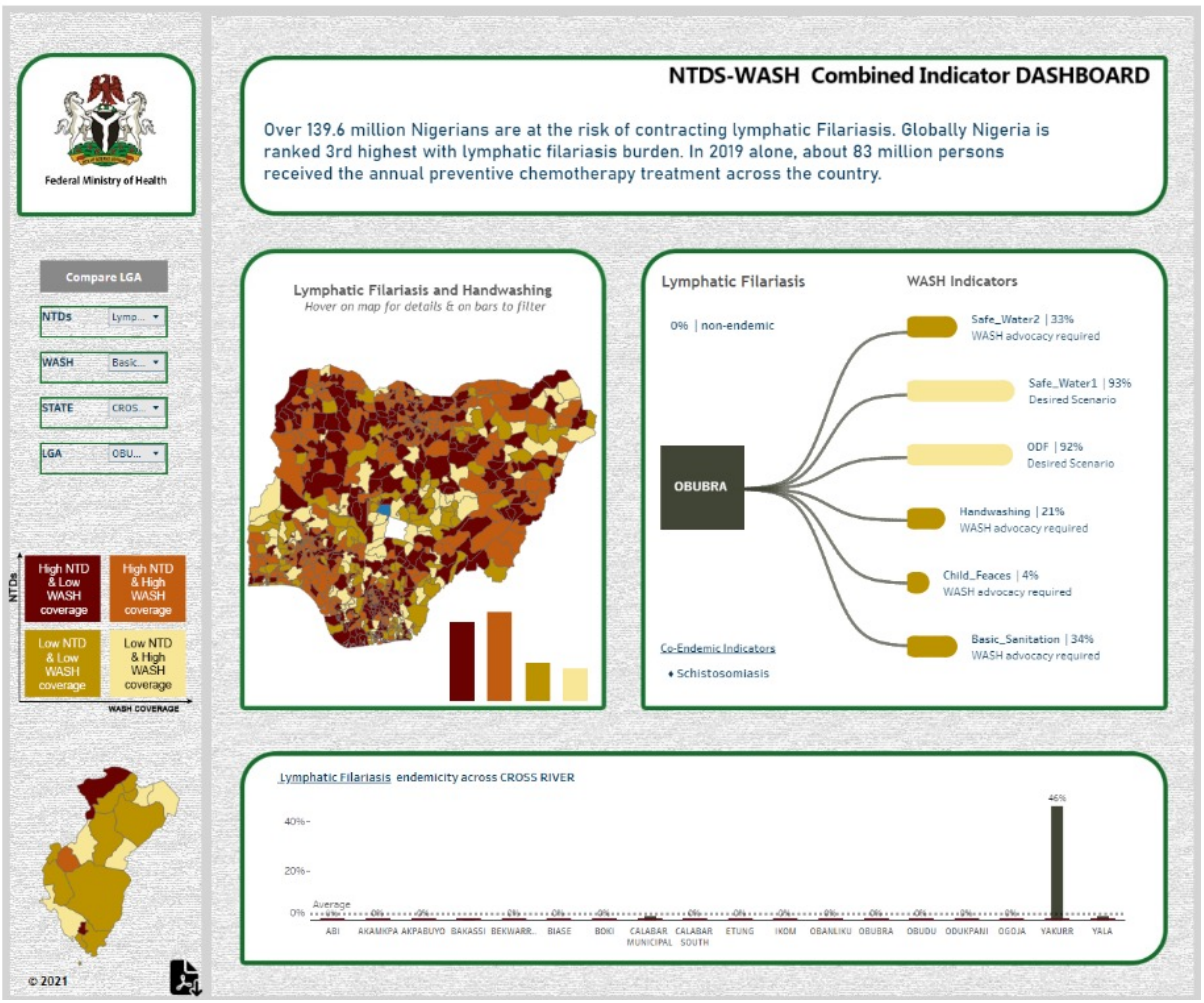
H/L WASH	H/L NTD	2x2 table
1	1	=F3&G3
0	1	01
1	0	10
0	1	01
1	0	10
1	1	11
0	0	00
0	0	00
1	0	10
1	1	11
0	0	00
1	0	10
1	0	10
1	0	10
0	0	00
0	0	00
0	1	01
0	1	01



Example of WASH and NTD data merge: Nigeria

In 2019, Sightsavers conducted an initial situation analysis to identify existing national systems for both WASH and NTD data. For WASH, the Department of Water Quality Control and Sanitation, part of the Federal Ministry of Water Resources (FMWR), developed a database for water and sanitation monitoring and evaluation data called WASH Information Management System (WASHIMS). For NTD data, there is no single national NTD database. NTD disease prevalence data were sourced from the Joint Reporting Form (WHO/UNICEF) (available on the ESPEN Platform: https://espen.afro.who.int/countries/nigeria).

Fig. 1. **WASHIMS data visualization**



The NTD national programme worked with the FMWR to develop a combined dashboard for WASH and NTD data. The WASH NTD data merge in Nigeria is hosted by WASHIMS and uses Tableau software for data visualization (Fig. 1). This tool is used by WASH and NTD technical working groups and other key stakeholders to explore the relationship between NTD prevalence data and WASH data elements in order to direct the appropriate interventions.