GlobalMix Mozambique Aim 2 participant data extraction.

MC Kiti

We use this file to extract participant metadata for further analysis.

Summary of participant characteristics by site

Table 1 shows the total number of individuals who were recruited into the study and their demographic characteristics by site.

Table 1: Particiant data from Mozambique sites

Variable	**Rural**, N = 415	**Urban**, N = 310
isindex	69 (75%)	66 (67%)
Unknown	323	211
participant_sex		
Female	244 (59%)	181 (59%)
Male	170 (41%)	128 (41%)
Unknown	1	1
participant_age		
<6mo	43 (10%)	23 (7%)
6-11mo	27 (7%)	31 (10%)
1-4y	50 (12%)	43 (14%)
5-9y	64 (15%)	34 (11%)
10-14y	54 (13%)	31 (10%)
15-19y	40 (10%)	21 (7%)
20-29y	62 (15%)	60 (19%)
30-39y	34 (8%)	32 (10%)
40-59y	24 (6%)	25 (8%)
60+y	17 (4%)	10 (3%)
read_write	177 (43%)	173 (56%)

Variable	**Rural**, N = 415	**Urban**, N = 310
occupation		
Child	70 (20%)	54 (21%)
Unemployed	86 (25%)	39 (15%)
Student	111 (32%)	75 (29%)
Homemaker	6 (2%)	14 (5%)
Casual laboror	12 (3%)	12 (5%)
Farmer	33 (9%)	0 (0%)
Fisherman	1 (0%)	0 (0%)
Business person	6 (2%)	25 (10%)
Office worker	8 (2%)	32 (12%)
Retired	1 (0%)	3 (1%)
Other	14 (4%)	2 (1%)
Unknown	67	54
enrolled_school	140 (35%)	79 (27%)
Unknown	18	12

Rural sensor data

Here, we clean the sensor participant data. Steps are: 1. Check that all the sensor_ids are available. 2. Replace sensor_id written incorrectly. These were identified visually and via code. 3. Drop records with inconsistent sensor_id and share with data managers for checks. 4. Merge with hwid data. This was extracted from the sensors by data managers. 5. Check missing sensor_id and hwid data. Export records for DMs to check.

In total, there are 407 participants who were issued sensors from the rural site. Out of these, we have complete records for 287. This is after dropping records due to missing sensor_id and hwid.

> **Note**
>
> The hwid is an important identifier for the participants and is the primary key for the sensor data. Without the hwid, we cannot link the sensor data to the participant data collected in REDCap.

Table 2 shows the characteristics of individuals with missing sensor IDS and hwids (n=120).

Table 2: Rural records missing sensor_id and hwid data

Variable	N = 120
isindex	19 (86%)
Unknown	98
participant_sex	
Female	66 (55%)
Male	53 (45%)
Unknown	1
participant_age	
<6mo	14 (12%)
6-11mo	6 (5%)
1-4y	16 (13%)
5-9y	16 (13%)
10-14y	13 (11%)
15-19y	13 (11%)
20-29y	15 (12%)
30-39y	12 (10%)
40-59y	9 (8%)
60+y	6 (5%)

Urban sensor data

In total, there are 301 participants who were issued sensors from the urban site. Out of these, we have complete records for 243. This is after dropping records due to missing sensor_id and hwid.

Table 3 shows the characteristics of individuals with missing sensor IDS and hwids (n=120).

Table 3: Urban records missing sensor_id and hwid data

Variable	N = 58
isindex	11 (69%)
Unknown	42
participant_sex	
Female	32 (55%)
Male	26 (45%)
participant_age	
<6mo	2 (3%)
6-11mo	3 (5%)
1-4y	11 (19%)

Variable	**N = 58**
5-9y	3 (5%)
10-14y	3 (5%)
15-19y	4 (7%)
20-29y	14 (24%)
30-39y	7 (12%)
40-59y	9 (16%)
60+y	2 (3%)

The `echo: false` option disables the printing of code (only output is displayed).