Promoting Effective Data Visualization Among Reproductive, Maternal, Newborn, Child Health and Nutrition Decision-makers in Tanzania

June 2018
Acknowledgments

The National Evaluation Platform (NEP) is a rigorous new approach to compiling and analyzing health and nutrition data from diverse sources, so that the Government can get strategic, evidence-based answers to their most pressing RMNCH&N program and policy questions. This study was implemented as part of the NEP, with funding from Global Affairs Canada – grant number 7059904, and technical assistance provided by the IIP-JHU.

The authors would like to acknowledge Hereswida Zeswida Monyiechi for her involvement during data collection, transcription, and translation support. We would also like to thank Khatib Khatib, Arnold Mwijage, and Shabani Said for their assistance with data collection, Ashley Sheffel for her technical assistance, and Steve Gesuale, Amanda Makulec, and Jessica Pomerantz for their initial input on the study design. Finally, we would like to thank NEP Technical Task Team (TTT) and High-Level Advisory Committee members that helped support this study.
# Table of Contents

- Introduction ........................................................................................................................... 4
- Study Aim and Research Questions ....................................................................................... 5
- Methods ...................................................................................................................................... 5
- Results ....................................................................................................................................... 7
- Conclusions ............................................................................................................................. 23
- Recommendations .................................................................................................................. 24
- References ................................................................................................................................. 24
Creating digestible ways to portray data is key to facilitating the use of data for decision-making. Although a wealth of data on reproductive, maternal, newborn, child health, and nutrition (RMNCH&N) exists in Tanzania and other countries in Africa, data are not consistently utilized to inform RMNCH&N program and policy decisions. National statistical offices and M&E departments in these contexts are engaged in data collection and analysis. However, there is an incomplete understanding of how to effectively share data in ways that RMNCH&N stakeholders can easily understand and apply to policy decisions. Policymakers vary widely in their formal training and experience with health and statistical content, and hold different job functions that influence how they use data. In practice, it’s difficult to distill key messages from data and explain statistical concepts like “uncertainty” with this audience. Understanding this audience’s needs is essential to determining more effective approaches for sharing data that encourages data-driven RMNCH&N policymaking.

There is substantial evidence on how data visualization—the process of graphically displaying data to tell a story—impacts the way individuals interpret data. There are two purposes to data visualization: communicating a story and sense-making. Given that human brains more rapidly process visual information compared to text, data visualization principles emphasize optimizing visuals that are attune with humans’ short-term memory (1-3).

In the health sector, research has explored how different verbal communication methods can influence patient understanding and decision-making, but has not examined interpretation of specific graphic types among audiences (4). For example, over the past thirty years, geographic information systems (GIS) technology has become a powerful tool to create maps that illustrate population-based health trends (5-7). However, there are no known studies on how maps that display health data are interpreted by the public or policymakers.

There are limited studies of the impact of data visualization on public health policymaking and data visualization interpretation in low- and middle-income countries (LMICs) (8). Most research on data visualization is in fields other than public health, such as business, and are focused on well-resourced settings (9). MEASURE Evaluation conducted a study to explore how different types of data visualization approaches are used to inform HIV programs in Namibia, South Africa, Tanzania, and Zambia (10). MEASURE Evaluation exclusively interviewed President’s Emergency Plan for AIDS Relief (PEPFAR) implementing partners. MEASURE Evaluation’s report highlighted the need to engage data users with developing visualizations as well as the need for training to interpret visualizations. As a separate activity, MEASURE Evaluation and Matchboxology, a human-centered design firm, conducted exploratory interviews and workshops with district-level HIV data users in Tanzania and South Africa to identify facilitators and barriers to using HIV data for decision-making. Human-centered design is a problem solving process that relies on end user perspectives to develop solutions (11). The MEASURE Evaluation and Matchboxology activity focused on using human-centered design to ultimately identify recommendations and prototypes to encourage HIV data use at the district-level. Prototypes proposed to address data use barriers in Iringa and Dar es Salaam included data use scorecards, health facility feedback reports, and mechanisms to improve data awareness and supervision. All districts in the study reached a consensus that a successful health information system (HIS) would offer “clear, easily understandable, and user-friendly information that empowers every data user to make informed and effective program decision making.” (12)
Study Rationale

This study explored how RMNCH&N program implementers and policymakers (“RMNCH&N decision-makers”) in Tanzania interpret and prefer different ways to visualize RMNCH&N data to inform how to better visualize and communicate data for RMNCH&N policymaking.

We conducted this study as part of the National Evaluation Platform (NEP)’s evaluation of the Government of Tanzania’s One Plan for Maternal Child Deaths in Tanzania (2008-2015) (“One Plan”). The NEP is a rigorous new approach to compiling and analyzing health and nutrition data from diverse sources so that national policymakers in LMICs can evaluate the effectiveness and impact of RMNCH&N programs and policies. NEP aims to bridge the gap between findings and policymaking. Based on interest from the Ministry of Health, Community Development, Gender, Elderly and Children (MOHCDGEC), NEP undertook an evaluation of the One Plan’s maternal components. NEP conducted the data visualization study to support dissemination of the One Plan evaluation and help develop best practices for translating data to RMNCH&N policymakers in Tanzania. This is the first time a study has been conducted on the relationship between data visualization and RMNCH&N policymaking in any LMIC.

Study Aim and Research Questions

Study Aim

The aim of this study is to characterize data visualization interpretation capacity and preferences among RMNCH&N decision-makers in Tanzania.

Research Questions

1. How do RMNCH&N decision-makers interpret RMNCH&N data?
2. What are RMNCH&N decision-makers’ preferences on visualizing and communicating RMNCH&N data? How do these preferences compare to data visualization “best practices”?
3. How do different types of RMNCH&N data visualizations influence RMNCH&N program and policymaking in Tanzania?

Methods

Data Collection

We conducted a total of 25 semi-structured in-depth interviews (IDI) with RMNCH&N decision-makers in different departments of the Ministry of Health, Community Development, Gender, Elderly and Children (MOHCDGEC), President Office – Regional Administration and Local Government (PORALG), and Tanzania Food and Nutrition Centre (TFNC). We focused on these government institutions given their role in RMNCH&N decision-making in Tanzania. We used a snowball sampling strategy to identify individuals to interview. We received initial suggestions through NEP Technical Task Team (TTT) members and ministry departments. We identified additional individuals to interview through recommendations from initial respondents. We also aimed to interview individuals that represented a range of seniority, specialty, and gender.

The NEP Tanzania team based at the National Bureau of Statistics, NEP Tanzania Resident Advisor, and an independent Tanzania-based qualitative researcher conducted all interviews in Kiswahili. Three student interns from the Eastern Africa Statistical Training Center (EASTC) served as supplemental note takers. Interviews took place in Dar es Salaam and Dodoma. A
A team of two individuals facilitated each interview— one lead interviewer and one note taker. A student intern joined most interviews as an additional note taker.

Each IDI explored respondents’ background and knowledge of statistics, and data visualization experiences. The IDIs also included three activities to help the research team understand how respondents interpret RMNCH&N data and respondents’ preferences for different types of data visualizations. We used only used graphs that could be made using Microsoft Excel or the R statistical package. The study omitted complex infographics. See Figure 1 for an overview of the IDI structure, description of activities, and links to research questions.

**Activity 1 – Explain key messages for 5 prepared graphs**

**Aim:** Understand decision-makers capacity to interpret RMNCH&N data.

**Description:** Interviewers showed participants five different visualizations and asked participants to identify key messages. Interviewers presented each graph one at a time and did not provide any assistance to participants with interpreting the graph.

**Activity 2 – Rank different data visualizations based on key message & Activity 3 - Rank different data visualizations that display specific statistical concepts based on key message**

**Aim:** Identify decision-makers’ preferences among different data visualization approaches.

**Description:** During both Activity 2 and 3, interviewers presented participants three different sets of cards. Each set of cards included one key message and different data visualizations based on the same data. Interviewers asked participants to rank the cards for each set based on which most clearly communicated the key message provided. Activities 2 and 3 included an assortment of graphs (bar graphs, dot plots, slope graphs, line graphs, pie charts, line graphs). Although structured similarly, Activities 2 and 3 differed slightly as Activity 3 included statistical concepts (“statistical significance”, proportion) in key messages.

---

**Research Question 1 (RQ 1): How do RMNCH&N decision-makers interpret RMNCH&N data?**

**Research Question 2 (RQ 2): What are RMNCH&N decision-makers’ preferences on visualizing and communicating RMNCH&N data?**

**Research Question 3 (RQ 3): How do different types of RMNCH&N data visualizations influence RMNCH&N program and policymaking in Tanzania?**
An independent qualitative researcher, who was part of the interview team, transcribed and translated all IDI recordings. Other members of the interview team cross-validated a sample of the transcripts to confirm accuracy.

**Ethics**

The National Institute for Medical Research of Tanzania granted IRB approval for the study in March 2017. In December 2016, Johns Hopkins Bloomberg School of Public Health (JHSPH) deemed the study non-human subjects research as the key-informant interviews collected no personal or private information. All US-based JHSPH researchers received research permits from the Tanzania Commission for Science and Technology in May 2017. Each respondent provided verbal and written consent to participate in the study.

**Data analysis**

We analyzed all 25 transcripts using the qualitative analysis software Dedoose. We included all transcripts in the analysis. We generated codes based on themes that arose in the IDIs. We used a grounded theory approach to review data and identify findings.

**Results**

**Participant Demographics**

We interviewed 25 individuals that specialize in overarching national health strategy, immunization, nutrition, malaria, and reproductive and child health. Participants represented a mix of mostly senior and mid-level professionals. We defined seniority based on position title: “Senior” = Principles and Directors, “Mid-level” = Senior Officers and Program Officers; “Junior” = Officers. See Figure 2 for a summary of participant characteristics.

![Professional Experience](image)

**Data is a critical aspect to everyone’s job**

All participants emphasized that using data is a key aspect of their day-to-day responsibilities for improving programs and policymaking. Data is used for monitoring & evaluation (M&E) of program and policy performance, advocacy of public health issues, commodity forecasts, budget justification, performance-based awards, and reports to international accountability.
frameworks (e.g. Family Planning 2020, Sustainable Development Goals). Participants shared that data highlighting poor performance is used to identify where they need to intervene – including targeting regions that need extra attention. Data is also used to advise partners on investment priorities.

“Data are the most important in my current role because you can use data to make a politician or policymaker understand a problem…data is the strength to explain any problem.” – Senior, Nutrition

Despite the role that data plays in the everyday work of participants, many participants reported having limited formal training opportunities in statistics or data use since graduating from university. Several participants shared that they have never received any on-the-job training in statistics. Participants that have received training attended workshops on M&E, DHIS 2, Demographic Health Surveys (DHS), and STATA. A few participants shared that they receive informal training and mentorship in data use from colleagues.

“Some say…‘send us to training’…how can I send a person to training while I was never trained on data?” – Mid-level, RCH

Capacity to interpret graphs is mixed

To gain a sense of participants’ capacity to interpret data, interviewers provided participants with five different graphs and asked participants to identify key messages for each graph (Activity 1). Interviewers did not provide any assistance to participants with interpreting graphs and did not disclose key messages identified by the study team. Below is a summary of findings for each graph.


- 1999
- 2004
- 2010
- 2015

Activity 1 - Card 1
Key messages identified by participants | Key messages from study team
--- | ---
Most participants identified that exclusive breastfeeding increased between 1999 and 2015. Several participants commented about changes in predominant and partial breastfeeding. A few participants mentioned that infants not breastfed has decreased or remained constant over time. Many participants tried to estimate percentages of breastfeeding coverage. Two participants did not describe any trends in the graph and repeated the title of the graph as the key message. | • Increase in exclusive breastfeeding between 1999 and 2015.  
• Proportion of infants predominantly breastfed decreased between 1999 and 2010 and increased between 2010 and 2015.  
• Proportion of infants not breastfed has remained constant.

"Exclusive breastfeeding in the year 1999 and 2015 is different - it is increasing. This means that women’s understanding on the important of breastfeeding is increasing as time goes. Those who were doing predominant are decreasing, which means there is something to look at. Those in partial to some extent also are decreasing. But what gives joy is that those in none are also decreasing more." – Senior, General

Graph 1 Summary

<table>
<thead>
<tr>
<th>Year</th>
<th>ANC4+ coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>59.1%</td>
</tr>
<tr>
<td>2010</td>
<td>39.7%</td>
</tr>
<tr>
<td>2015</td>
<td>49.2%</td>
</tr>
</tbody>
</table>

Graph 2 Summary

<table>
<thead>
<tr>
<th>Year</th>
<th>ANC4+ coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>59.1%</td>
</tr>
<tr>
<td>2010</td>
<td>39.7%</td>
</tr>
<tr>
<td>2015</td>
<td>49.2%</td>
</tr>
</tbody>
</table>

Key messages identified by participants | Key messages from study team
--- | ---
Nearly all participants mentioned that ANC coverage decreased between 2004 and 2010 and increased between 2010 and 2015. Less than half of participants mentioned that ANC coverage in 2015 did not reach the One Plan Target of 90%. A few participants suggested reasons for the decrease and increase in ANC coverage. | • ANC coverage decreased between 2004 and 2010 and increased between 2010 and 2015.  
• Tanzania Mainland did not reach the One Plan target for ANC coverage.
Additional under-five lives saved per year since 1999 in Mainland Tanzania due to intervention scale-up

The third graph showed to participants represented results from a Lives Saved Tool (LiST) (http://www.livessavedtool.org/) analysis, which identifies lives saved between 1999 and 2015 due to the scale-up of RMNCH&N interventions.

**Graph 3 Summary**

<table>
<thead>
<tr>
<th>Key messages identified by participants</th>
<th>Key messages from study team</th>
</tr>
</thead>
</table>
| Most participants struggled with interpreting this graph and felt overwhelmed by the volume of information and colors. Four participants did not describe any trends in the graph and repeated the title of the graph as the key message. Several participants refused to share any key message. Many participants did not mention lives saved and interpreted the graph as depicting increase coverage of interventions. | • Increase in lives saved between 1999 and 2015.  
• Malaria prevention & treatment, vaccines, and stunting and wasting interventions contributed the most to lives saved between 1999 and 2015. |

“It is very congested! What do I have to interpret here? I do not get a message here I just see it [as] confusing.” – Mid-level, RCH

“There is no key message here. It will bother me to read because separating these small colors. I am color blind. Let’s agree first that there is no key message here. There is many information cluttered in this single chart. It is telling me lives saved, but there is no key message here.” – Mid-level, Nutrition
Graph 4 Summary

<table>
<thead>
<tr>
<th>Key messages identified by participants</th>
<th>Key messages from study team</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half of participants did not comment on confidence intervals (CIs) or statistical significance depicted. Nearly all participants commented that maternal mortality decreased and then increased. Several participants identified how there has not been a statistically significant change in maternal mortality during this period.</td>
<td>• There was no statistically significant change in maternal mortality between 2004 and 2015.</td>
</tr>
</tbody>
</table>

Coverage of household insecticide treated nets/indoor residual spraying by region in Mainland Tanzania (TDHS 1999, 2010)
Graph 5 Summary

<table>
<thead>
<tr>
<th>Key messages identified by participants</th>
<th>Key messages from study team</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most participants identified that coverage of insecticide treated nets/indoor residual spray (ITN/IRS) increased among all regions between 1999 and 2010. Many participants shared trends they saw for specific regions. Several participants struggled with interpreting colors – possibly due to not understanding the key or color blindness. To aid in interpreting colors, participants suggested directly labeling regions on the map with ITN/IRS coverage instead of using colors, or assigning colors to ranges of ITN/IRS coverage within the key.</td>
<td>• ITN/IRS coverage increased dramatically for most regions between 1999 and 2010.</td>
</tr>
</tbody>
</table>

Data visualization choices are determined by many factors

Interviewers probed participants about how they learned to present data and factors considered when deciding what type of visualizations to use. Below are several factors that influence participants’ decision on what data visualizations to use.

**Audience**

Participants frequently cited audience as a key factor in deciding which type of data visualization to use. Data prepared by participants is shared with national and sub-national policymakers and program officers, donors, non-government organizations, health practitioners, academia, and the public. Participants are cognizant of the different education levels and motivations of these audience groups, which influence how they present data.

“Generally you look at your audience. Where do you want to present that data? The level of education is very important.” – Senior, Nutrition

**Simplicity and understandability**

Simplicity and understandability are overarching principles – participants strive to choose the simplest visualization that can be easily understood. Ensuring that a data visualization is understood by an audience is a high priority. Even though participants shared that they always choose the simplest way to present data, participants did not always agree on data visualization types considered “simple” and “understandable.”

“My intention is to make them understand, not give them an exam for them to fail. I use simple methods that I know they will understand at the end of the day.” – Mid-level, RCH

Most participants described simplicity and understandability in the context of the audience’s statistical capacity. Interviewers asked participants to describe data visualizations that they felt audiences could understand easier.
A majority of participants listed bar charts, pie charts, and maps as data visualizations that audiences tend to grasp easily, but participants also mentioned using pictures and words to convey data. Participants also shared that they sometimes use line graphs, icons, maps, and words for communicating data. Nearly all participants immediately described how they learned to present data in pie charts, bar charts, and tables. This list of graphs mirrors the types of graphs participants identified as being part of their training on how to present data.

“To a politician if you use bar chart, they can easily understand a bar that is long and short. Even with pie charts they can see rounds and segments and get a certain meeting.” – Senior, Vaccines

“If I am talking to people who are a bit educated, it is good to present through bar and pie charts as they do understand. For those who are less educated like common citizen, using words can be easier for them to understand than pie and bar charts. Telling common citizens and politicians deaths in absolute numbers rather than ratios is easier for them to understand.” – Senior, RCH

“Maps [are grasped easily] because policymakers want to know a snapshot of what is going on. When you project a map, they quickly will know where to invest.” – Mid-level, RCH

Participants shared inconsistent opinions on graphs that they consider difficult to understand. For example, participants shared conflicting opinions on whether audiences can easily understand tables. While some participants felt that tables are easy to understand, others felt that tables are more geared towards technical audiences and require the audience to interpret the data themselves. Depending on the audience, tables also prompt participants to dedicate more effort to explaining what tables are trying to show.

“The most difficult to understand are statistical tables. If you use those statistical data alone it is challenging for people to read. Because many of them have low understanding on statistical data...many of them are not taught data interpretation so it becomes very difficult. With graphs it becomes simple for them – ‘Ah so this means this.’” – Junior, General

Some participants also commented that less technical audiences have a challenging time understanding a graph that includes multiple indicators.

“Graphs are easy to present when they show data separately instead of combining [indicators]. Showing data combined confuses the audience and presenter.” – Junior, Nutrition
To help facilitate understanding, participants use formatting like specific colors or fonts. For example, some respondents described using red, yellow, and green to represent performance and the severity of a health issue. These colors can resonate with audiences of varying statistical capacities.

“You can present in black and white, but color can help people notice differences.” – Senior, Vaccines

“We use green to show areas that have done well and red to show regions with big problems that are not doing well. <The red regions> are in danger so if it is a matter of prioritizing, we have to focus on these areas. In most cases with maps, if something is serious you tend to use red to show danger so that you catch people’s attention. Those are colors that we mostly in presentations even in measuring malnutrition. Using red to show people most affected is alarming.” – Mid-level, Nutrition

Types of information to present
After identifying the audience, participants determine what data to present and how to present data based on audience interests and motivations. Participants mentioned preparing graphs based on what the audience will want to see or questions the audience may ask. For example, one respondent working in immunization shared that he prepares combination charts or pivot charts with coverage of several vaccine doses since he anticipates people viewing the chart will be interested in comparing coverage across vaccine doses.

“It depends on the audience. For example if I am reporting to my boss, he is very busy so he wants to see things quickly. For him I will make a graph with a caption. If I am writing a big report that needs to be presented somewhere, I mostly use tables and mix it up with graphs. The main criteria is the audience. If we are presenting to facilities, sometimes they may not understand graphs but they may understand tables because they work with many tables at the facility.” – Senior, General

A minority of participants shared that they choose a graph based on the key message they want to present. Participants mentioned the following:
- Pie chart to show proportion (e.g. distribution of disease and finances)
- Bar charts to show trends over time, compare performance among years or institutions
- Tables to show trends by region
- Maps to show trends by region

Interviewers also asked participants about their approach towards visualizing proportions/fractions and CIs. These concepts are of particular interest to the research team because they are critical to understanding NEP Tanzania’s results.

Participants shared that showing proportion can be difficult to some audiences – while some audiences want to see percentages, other audiences are only interested in absolute numbers. Respondents did not have a consensus on whether this preference is based on statistical background.
Participants were divided on whether depicting CIs is important, but many acknowledge that it depends on the statistical skills of the audience. However, when interviewers asked participants to define CIs, few participants could provide an accurate definition. Many participants also shared that policymakers have a limited understanding of CIs, and that because most people have not thought about CIs since university, they have a limited understanding of the concept. Some described CIs as an “academic concept.” Very rarely have participants seen CIs depicted in presentations.

“You know our people do not have time...you find a policymaker has <many> meetings so starting to tell them confidence interval stories...I think you will just be pouring water in the sack.” – Mid-level, Nutrition

**Capacity to explain the visualization**

A few participants commented that they choose a visualization based on their own capacity to present the visualization so that the audience understands.

“I choose a way which is easy for me to interpret the data. I can’t say that I would use a way that I am not experienced [with] or knowledgeable [about] so that I would fail to present the data.” – Mid-level, RCH

**Graph preferences**

Interviewers further assessed participants’ data visualization preferences through Activities 2 and 3. During these activities, interviewers gave participants sets of cards that included different data visualization options based on the same data and the same key message. Participants ranked the data visualizations in each set based on which most clearly showed the key message. Interviewers then asked participants to explain why they ranked the data visualizations in the order they specified, and how they could be improved.
**Activity 2 - Findings**

Set 1 Key Message: “ANC coverage varies by household wealth in Tanzania Mainland. The gap between the poorest and wealthiest has increased over time since 2004.”

---

**Participants’ response**

A majority of participants ranked the bar chart (graph one) the highest. Participants who chose graph one stated that they chose this graph because the bar graph is the easiest to understand. Many participants suggested that graph one could be further improved by orienting bars vertically rather than horizontally. Participants described graphs two and three as confusing.

**Data visualization research evidence**

A dot plot (graph two) is a preferred approach to visualizing equity because people can quickly and accurately interpret the space between dots on a common axis (1, 13). A dot plot also helps show how gaps in coverage among wealth quintiles increase and decrease over time. Dot plots are the standard graph used to represent equity by the University of Pelotas’ International Center for Equity in Health.
Set 2 Key Message: “About half of women receiving ANC in Tanzania Mainland reported being satisfied with how well a provider met a set of expectations about the services. Client satisfaction does not vary by type of facilities, but is higher in non-government facilities compared to that in government facilities.”

<table>
<thead>
<tr>
<th>Card 1</th>
<th>Card 2</th>
<th>Card 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranked 1st</td>
<td>92%</td>
<td>8%</td>
</tr>
<tr>
<td>Ranked 2nd</td>
<td>8%</td>
<td>54%</td>
</tr>
<tr>
<td>Ranked 3rd</td>
<td>0%</td>
<td>38%</td>
</tr>
</tbody>
</table>

Participants’ response
A majority of participants ranked the bar chart (graph one) the highest and ranked the dot plot (graph three) the lowest. One participant refused to rank cards 2 and 3 and stated that they were too complicated to understand. Participants described graph one as “straightforward” whereas graphs two and three looked “crowded.”

Data visualization research evidence
Bar graphs with more than 5 categories can be challenging to quickly interpret. A slope graph (graph two) is a preferred approach to compare two categories because people can quickly judge slope (1). However, since the percentages are very close, the slope graph is not necessarily ideal in this example.
Set 3 Key Message: “During the past 15 years, a majority of under-five deaths in Tanzania Mainland have been attributable to the same five causes: neonatal, measles, pneumonia, malaria, and diarrhea. Between 1999 and 2015, the proportion of under-five deaths attributable to AIDS has decreased, while the proportion of under-five deaths attributable to injuries has increased.”

Participants’ response
Participants ranked the pie chart (graph two) the highest and ranked the 100% stacked bar chart (graph one) the lowest. Participants were divided between the pie chart and bar chart (graph three) – both graphs that everyone stated they prefer to use throughout the study. Regarding graph one, participants found this more difficult to understand. One participant stated, “I wouldn’t recommend at all graph one. I will never recommend it or I wouldn’t use it so I will not talk about it at all. I can’t analyze it quickly.”

Data visualization research evidence
The 100% stacked bar (graph one) is a preferred approach to compare proportions from two different years. Studies show that people have difficulty accurately interpreting distance along curves and slice area, compared to 90 degree angles and straight lines (1). Accurately comparing pie charts is also a challenge (2). Since the key message focuses on proportion, a regular bar chart (graph three) should be the most challenging to interpret.
**Activity 3: Findings**

During Activity 3, interviewers showed participants different ways to visualize CIs and proportion. Activity 3 included three sets of cards that showed different ways to visualize CIs and proportions using the same data set. Similar to Activity 2, interviewers gave participants a key message for each data set and asked participants to rank cards within each set.

Set 1 Key Message: “Since 1999, contraceptive prevalence in Tanzania Mainland has increased. There is a statistically significant increase in contraception prevalence between 2004 and 2015.”

<table>
<thead>
<tr>
<th>Card 1</th>
<th>Card 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranked 1st 33%</td>
<td>67%</td>
</tr>
<tr>
<td>Ranked 2nd 67%</td>
<td>33%</td>
</tr>
</tbody>
</table>

**Participants’ response**

Participants preferred line CIs (graph two) over shaded CIs (graph one). Several participants felt there is no difference between the two graphs, yet chose a graph because the exercise forced them to choose. A couple of participants felt that the shading could be misleading. One participant commented that they had seen CIs depicted like in graph two, which is why they prefer this graph.

**Data visualization research evidence**

Research has demonstrated that error bars, like in graph two, can be inaccurately interpreted (14). Shaded confidence intervals are considered to be clearer. UNICEF has used shaded color blocks to represent 95% CIs in graphs depicting stunting among children under-five.
Set 2 Key Message: “ANC readiness—a measure of how well a facility can provide services—varies by region, however, the variation is not statistically significant.”

<table>
<thead>
<tr>
<th>Card 1</th>
<th>Card 2</th>
<th>Card 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranked 1st</td>
<td>60%</td>
<td>21%</td>
</tr>
<tr>
<td>Ranked 2nd</td>
<td>24%</td>
<td>50%</td>
</tr>
<tr>
<td>Ranked 3rd</td>
<td>16%</td>
<td>29%</td>
</tr>
</tbody>
</table>

Participants’ response
Participants ranked the bar chart with CIs as lines (graph one) first and the dot plot with CIs as shaded bars (graph three) last. Participants shared that they preferred graph one the most because it is a bar chart. One participant did not rank graphs two and three, and claimed these graphs are identical.

Data visualization research evidence
Based on the evidence of error bars being potentially misleading and the number of regions depicted, graph three is a preferred approach towards depicting CIs.
Set 3 Key Message: “In Tanzania Mainland, about half a million of the under-five deaths during the past 15 years are attributable to stunting or wasting. Across all zones, stunting and wasting contributed to a smaller proportion of total under-five deaths compared to that of other causes.”

Participants’ response

Participants ranked the 100% stacked bar (graph 2) over the regular bar chart (graph 1). Based on the interviewer’s questioning on why participants prioritized each graph, participants that chose graph 2 could have not read the key message carefully and prioritized graph 2 because they preferred seeing percentages rather than absolute numbers.

Data visualization research evidence

Of the two graphs, only the first graph would allow someone to determine the number of under-five deaths due to stunting and wasting, which was a part of the key message.
Challenges

Participants shared several challenges to visualizing and communicating RMNCH&N data. Challenges are listed in order from most to least frequently mentioned.

**Statistical and presentation skills of policymakers**

A majority of participants shared that the biggest challenge is limited statistical capacity of decision-makers. This includes both policymakers who receive and present data. As mentioned before, most policymakers have not had training in data literacy or statistics since university. Participants shared that policymakers they present to struggle with interpreting data and are resistant to “statistical jargon” and “complicated data.” This also includes identifying a key message from the information they receive.

“The greatest existing challenge that I see is that some policymakers do not have the knowledge to interpret or present data so it creates controversy in decision making. You can meet a decision maker who gives a statement that jeopardizes people, and it has some influence because of the popularity of that person. However, that person did not give consideration to the data and its meaning, so a decision maker’s understanding is sometimes an issue. This means we have to do extra work in data presentation – how do we make our policymakers and decision makers translate data before making decisions.” – Senior, Nutrition

Similarly, participants expressed that as presenters, there is a level of skill needed to create an appropriate visualization. Participants shared that decision-makers presenting must have enough statistical knowledge to be able to explain data to the audience, and answer any questions from the audience. Participants additionally mentioned that figuring out how to present to different audiences, and most importantly, how to get the audience to care is a consistent challenge.

“The challenge is that you must understand data analysis so you can present to an audience with different levels of understanding. You can start presenting your percentages and everyone is sitting there with no questions. Higher-level people cannot tell you ‘I do not understand you.’” – Mid-level, RCH

**Trust in data**

Several participants shared that the audience can question the validity of data presented, which can compromise policymakers’ trust and use of data. This happens when decision-makers do not understand how data was collected or data source. Another concern is that policymakers may not want to accept data they receive if the data is not favorable. One participant shared a specific example of this relating to the DHS 2015/16 revealing an increase in maternal mortality. Policymakers expressed reluctance to accept that maternal mortality increased, which prompted discussion on what should be considered the “true” maternal mortality ratio.

“Another problem that I see is that people don’t believe in statistical data. You can present data, but you find a leader or politician saying this data is not right! The success of politicians [can be] based on data quality” – Senior, RCH
External factors affect translating RMNCH&N data to policy

Even when equipped with provoking data, participants cited many other factors—indeed, independent of the visualization and presentation approach—that influence whether RMNCH&N data can be translated into policy. This includes whether there is sufficient funding to support policy changes proposed, existence of political will and interest, human resources needed to implement policies. A few participants mentioned how data presented must be tied to specific policy timelines to help ensure that findings have an impact.

Suggestions

Participants provided many suggestions regarding best practices when visualizing and communicating RMNCH&N data and how to improve data visualizations. A majority of participants requested training on basic data literacy, statistics, data visualization, and data presentation to improve their capacity to receive and apply data. Figure 3 lists a summary of suggestions organized by theme.

| Data visualization (format & use) | • Use vertical bar graphs rather than horizontal bar graphs. |
| • Label values directly on graphs. |
| • Include a key with all graphs. |
| • Write a short description of trends (a key message) adjacent to graphs. |
| • Include grid line backgrounds on graphs. |
| • Use “simplest” graph possible to visualize data. |
| • Avoid including multiple indicators within a single graph. |
| Color | • Use colors that quickly highlight issues (red and green) and a color palette that has sufficient distinct, bold colors. |
| • Use color-blind friendly formatting. |
| • Limit total number of colors in graphs. |
| Presentation | • Prepare more abbreviated materials rather than long reports. |
| • Ensure that messages are understandable and are written using simple language. |
| Training | • Develop curriculum and host trainings on basic data literacy, statistics, data visualization, and data presentation for policymakers. |
| • Provide orientation to explain how NBS collects data and calculates indicators. |

Figure 3. Data visualization suggestions

Conclusions

Study participants universally believe that data is critical to their roles improving RMNCH&N outcomes in Tanzania. Participants agree that striving towards an environment where applying data to policymaking is ideal. However, participants acknowledge that equipping policymakers with relevant RMNCH&N data does not necessarily translate to data-informed policies.

This study suggests that data literacy and statistical skills among Tanzanian decision-makers is variable. Most participants demonstrated awareness of many critical factors that should influence a visualization choice—audience, key message, simplicity—but assessments of data interpretation and preferences suggest that data literacy may be weak. Many decision-makers may be relying on training in data visualization that they received many years ago in university. There also appears to be some discomfort with interpreting and using graphs that are not bar charts, pie charts, and maps. Candid about the limitations of their statistical
skills, decision-makers expressed enthusiasm for expanding data interpretation and presentation skills. Addressing gaps in basic data literacy and presentation skills among policymakers is vital to bridge the gap between data and policymaking.

**Recommendations**

Host basic data literacy, statistics, and data visualization training for government officials supporting RMNCH&N in Tanzania.

- Training opportunities were the most common suggestion among participants. Given many haven’t received any training since university—sometimes decades ago—participants felt that it would be highly beneficial to have refresher trainings on statistics and data interpretation. Furthermore, although participants stated and demonstrated a preference for bar charts, pie charts, and maps, there is an eagerness to learn more about different types of data visualizations and how to create simple data visualizations with impact. Bar charts and pie charts are both graphs also mentioned as the first types of graphs participants learned using, and perhaps continued preference for these graphs reflect a lack of training since university.

Expand skills in developing key messages for RMNCH&N decision-makers.

- Many participants struggled with identifying key messages when shown a graph and stated that a major challenge is figuring out how to present data to get people to care. This encompasses both having the statistical expertise to confidently present data, but also an ability to refine presentations based on the statistical capacity of an audience.

Develop data visualization standards for presenting RMNCH&N data.

- Study findings demonstrate that decision-makers have specific preferences on the types and formatting of graphs. This is particularly important in terms of color; many participants stated that they are color-blind or have colleagues that are color-blind. Only one participant stated that their department had data visualization and presentation standards.

**References**

2. Evergreen SD. Effective data visualization: The right chart for the right data. SAGE Publications; 2016 May 3.