

Review

Not peer-reviewed version

---

# Towards a Common Understanding of Gender-Responsive Monitoring and Evaluation for Health Programs and Interventions: Evidence from a Scoping Review

---

[Anna Kalbarczyk](#)<sup>\*</sup>, Daniel Krugman, Shatha Elnakib, Elizabeth Hazel, Amy Luo, [Rosemary Morgan](#)

Posted Date: 24 May 2024

doi: 10.20944/preprints202405.1597.v1

Keywords: gender; gender responsive; monitoring and evaluation; gender equity



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

# Towards a Common Understanding of Gender-Responsive Monitoring and Evaluation for Health Programs and Interventions: Evidence from a Scoping Review

Anna Kalbarczyk <sup>1,\*</sup>, Daniel Krugman <sup>1</sup>, Shatha Elnakib <sup>1</sup>, Elizabeth Hazel <sup>1</sup>, Amy Luo <sup>2</sup> and Rosemary Morgan <sup>1</sup>

<sup>1</sup> Department of International Health, Bloomberg School of Public Health

Baltimore, MD USA; akalbarc@jhu.edu; dkrugma1@jhu.edu; selnak1@jhu.edu; ehazel1@jhu.edu; rosemary.morgan@jhu.edu

<sup>2</sup> Department of Population, Family and Reproductive Health, Bloomberg School of Public Health, Baltimore, MD, USA; aluo9@jhmi.edu

**Abstract:** Given the many approaches to and definitions of gender responsive monitoring and evaluation (M&E) for health programs and interventions there is a lack of clarity on how to operationalize it including what to measure and how to measure it. We conducted a scoping review to understand what makes M&E gender responsive. We included 31 studies and conducted two rounds of extraction to delineate ways in which gender was integrated into M&E. Twelve articles described the use of theory to guide M&E though most were not related to gender. Twelve articles employed a gender score in data collection, most of which measured Likert scale responses related to gender equity. Even though most studies did not use a specific gender framework, most incorporated gender domains in their analysis. Seven studies used participatory methods in the design and implementation of M&E. Most studies conducted M&E on programs or interventions that were designed to be gender intentional and related to gender issues. Gender responsive M&E intentionally integrates gender into the M&E process, regardless of how gender-intentional the program or intervention is. Gender dimensions can be identified through gender theories, models, scores, and frameworks to inform tool development, data collection, analysis, and stakeholder engagement processes.

**Keywords:** gender; gender responsive; monitoring and evaluation; gender equity

## Introduction

There are different approaches to, and definitions of, gender responsive monitoring and evaluation (M&E) in the context of health programs and interventions. Other terminology used to describe similar and related concepts includes gender-intentional, gender-sensitive, and gender-specific M&E, among others. The lack of common understanding is unsurprising given diverse understandings of gender, which itself is multi-faceted, including both gender identity and gender power relations. Gender identity refers to whether a person identifies with being a woman, man, or gender minority individual [1]. Gender power relations explore the ways in which gender inequities manifest through interpersonal, communal or societal relations, such as through inequitable access to resources, roles and practices, norms and beliefs, decision-making power and autonomy, and laws, policies, and institutions [2]. Gender is also often incorrectly conflated with biological sex or equated solely with women and girls.

The complexity of gender is compounded by its intersection with other social stratifiers, such as race and age, its context and time varying nature (i.e., how gender power relations manifest in one context may not be the same as in another context and understanding of gender and gender power relations change over time), and its embeddedness within the structures and institutions that encompass health. These definitional complexities lead to complexities in its measurement, including

difficulty in teasing out relevant factors that may have an impact on a specific health program or intervention.

M&E is the process of tracking progress and determining effectiveness of a program or intervention. It is important for defining program priorities, resources, and activities and helps inform decision-making and learning, as well as establish accountability. Integrating gender into programs and interventions ensures that the ways in which underlying gender inequities might positively and negatively affect program implementation and outcomes are considered and addressed which helps to ensure program outcomes are achieved, the needs and priorities of specific groups (particularly those that are more marginalized) are met, and programs and interventions do not perpetuate harmful gender inequities.

Operationalizing gender responsive M&E has been variable and applications often include different components. How gender responsive M&E is operationalized is important, as it affects what actions are taken. For example, within the literature, gender responsive M&E has been described as: 1. measuring underlying gender inequality that leads to poor health outcomes and inequities between and within groups (usually between men and women) [3]; 2. whether a program or intervention addresses the different priorities and needs of men and women [4]; 3. whether a program or intervention addresses gender inequities to improve health and wellbeing; and/or 4. measuring and evaluating differential effects of an intervention or program on groups (usually men and women) as well as any unintended consequences that may disadvantage one group over the other [5]. Some definitions include process-related considerations, including how and by whom data is collected and analyzed, and/or whether a M&E process is inclusive and participatory of diverse voices, particularly women's voices [5]. Many definitions include a combination of the above.

Consequentially, within gender responsive M&E, there are different conceptualizations of what indicators should be used to track progress and determine effectiveness. Types of indicators range in types, including sex-specific (focusing on women or men or sub-groups of women and men), sex-disaggregated (comparing women or men), or gender equality indicators (underlying gender inequality that leads to poor health outcomes and inequities) [3]. Some view the inclusion of sex/gender disaggregated data or indicators as sufficient to be considered gender responsive, while others think that gender M&E indicators need to go beyond sex/gender disaggregation to measure gender inequality and progress towards gender equality [6]. The overall lack of consensus on and operationalization of gender responsive M&E indicators has stymied clarity and integration of gender within programs and interventions [7].

Considering these differing approaches, we conducted an exploratory scoping review of the academic peer reviewed literature to examine the nature of how gender has been incorporated into the M&E of health programs and interventions, with a particular focus on reproductive, maternal, newborn, children and adolescent health, and nutrition (RMNCAHN). The review was conducted as part of the Monitoring and Action for Gender and Equity (MAGE) project which aims to galvanize gender responsive M&E for RMNCAHN, which is why we choose to focus on RMNCAHN. The aim of this review is to develop a consolidated evidence base of how gender has been integrated into M&E for RMNCAHN programs and interventions to better inform gender responsive M&E, with the understanding that different (and often inconsistent) terminology and approaches have been used. Ultimately, we wanted to answer the question: what makes M&E gender responsive within RMNCAHN programs?

## Methods

We followed the protocol defined by the Preferred Reporting Items for Systematic Review and Meta-Analysis extension for Scoping Review (PRISMA-ScR). This review was not registered in PROSPERO because the platform does not accept registrations for scoping review protocols.

We conducted a scoping review of peer reviewed publications in 3 databases: PubMed, CINAHL, and SCOPUS. The search was developed by creating search terms that corresponded to three concepts, "gender", "monitoring and evaluation", and "RMNCAHN". Search terms were developed in consultation with a global health informationist and faculty with expertise in each of

these three areas. The search was conducted on May 27, 2022. The final search strategy is provided in Supplemental File 1.

Articles were included if they were in English, published between 2010-2022, focused on RMNCAH topics, described a M&E component, and incorporated gender. Articles were excluded if they did not meet all inclusion criteria.

### *Screening Process*

After removing duplications, references were imported into the Covidence reference manager platform. Two independent reviewers conducted the title and abstract screening and subsequent full-text review; conflicts were resolved by a third reviewer.

### *Data Extraction and Analysis*

Given the novelty of this review and the diverse array of articles that were included under the selection criteria, we developed a two-review extraction process. In the primary extraction, one author (DK) used the extraction tool in Covidence to broadly extricate the different ways the included articles incorporated gender into their M&E processes. Through this, we were able to get a sense of the different ways in which gender was integrated into the M&E of RMNCAH programs and interventions to gain a better understanding of what gender responsiveness in the context of M&E is. Due to the lack of a common approach to gender responsive M&E, however, many authors did not refer to their M&E processes as being gender responsive or explain how gender was integrated into their M&E. With this perspective, two authors (RM and AK) then conducted a second round of extraction and exclusion based on emerging and relevant themes in an Excel matrix, delineating different ways in which gender was integrated based on our knowledge of gender integration. The cumulative results from these two extractions are featured in the results section.

## **Results**

### *Search Results*

We identified 4555 articles which were imported into Endnote for deduplication; 896 duplicates were removed. We screened 3659 references and of those, 3595 references did not meet inclusion criteria. Sixty-four full-texts were then reviewed by two independent reviewers; 27 studies were excluded for not having a gender focus (n=17), no available full-text (n=5), protocol only (n=4), and no M&E (n=2). In the second round of extraction, six more articles were eliminated because they didn't meet the inclusion criteria. Thirty-one articles were included in the final analysis. The PRISMA diagram (Figure 1) depicts the selection flow diagram.

### *Overall Description of Included Studies*

Supplemental File 2 outlines the articles included in the extraction. Of the 31 articles that met the inclusion criteria, 26 (84%) were direct evaluations of interventions [8–33], four (13%) were systematic literature reviews of gender focused methods and evaluation tools [34–37], and one was a review of a large multi-site program [38]. Fourteen of the 26 (45%) evaluations were conducted in Sub-Saharan Africa [9–12,16,21–24,26,28,31–33] while the remaining articles focused on specific locations were split between North America (n=4) [17,18,27,29], the Middle East (n=4) [8,13,15,30], and India (n=4) [14,19,20,25].

The most prevalent RMNCAH focus area was sexual health (n=16, 52%) [13,18–23,25–28,30,32,33,37,38], largely concerning the issues of gender-based violence (GBV) and intimate partner violence (IPV) (n=8) [13,19,23,26,30,33,35,37] and HIV/AIDS prevention (n=8) [19,22,25,28,32,33,37,38]. Seven articles focused on maternal and neonatal health interventions [9,11,12,15,16,26,29], ten evaluated adolescent or child health and nutrition programs [8,12,15–18,20,22,27,31], and three assessed family planning-focused programs [24,32,36].

### *Types of Studies*

Most studies employed quantitative methods in the evaluation. Sixteen of 31 studies (52%) were purely quantitative [8–13,17–19,21,23,25,28,29,31,32], while an additional six studies used qualitative methods as part of a mixed-methods design [14,15,20,22,30,38]. Surveys were the most common instrument used to collect data through randomized control trials (n=5) [8,12,18,21,28], cluster randomized control trials (n=6) [9,19,23,25,29,31], cross-sectional cohort studies (n=3) [11,17,32], and controlled cohort studies (n=1) [10]. In the eight mixed-methods and six qualitative-only evaluations [16,21,24,26,27,33], the most popular methods employed were in-depth interviewing (n=7) [14,16,20,22,26,27,33], focus group discussions (n=4) [16,20,22,26], and semi-structured interviewing (n=3) [15,24,26]. However, two articles employed unique methods surrounding ethnography [26,27]. Lees et al. utilized hearsay ethnographies where participants chronicled their daily lives in diaries after receiving training on ethnographic observation [26], while Exner-Cortens et al. employed Photovoice, a qualitative research methodology where participants use photography to capture their experiences around a health intervention [27].

Four articles were systematic reviews of the existing literature [34–37]. Each had different focuses. Adamou et al. explored gaps in male engagement in family planning programs (Adamou et al., 2019). Sharma et al. reviewed “promising practices” for the monitoring of gender-based violence interventions in humanitarian settings around the world [35]. More broadly, Mandal et al. reviewed measures of women’s empowerment and other gender-related constructs in family planning and maternal health interventions [37], while Kowalczyk et al. broadly surveyed issues of conducting and evaluating gender-based programs in global health [34]. Adamou et al. and Sharma et al. also employed in-depth interviews with their reviews to give more depth to their findings [35,36].

### *Ways in Which Gender Was Integrated into M&E*

Through the data extraction process, we mapped the existing evidence on gender responsiveness in M&E. These included: the use of gender-related theory or frameworks to guide M&E, the inclusion of gender scores, the inclusion of gender domains of analysis, the disaggregation of data by sex, and the use of community-based/participatory approaches.

#### *Use of gender-related theory, models, or frameworks to guide M&E*

Many tools and frameworks used in gender integration are grounded in theory, both gender theory and theory that is relevant for understanding gender. Gender theories include, for example, the theory of gender and power [39] and male gender norms theory, such as masculinities and hegemonic masculinities [40]. Other cross-cutting theories were used to help contextualize gender as a construct and the ways it manifests, such as intersectionality, the socio-ecological model, life course approach, and social cognitive theory.

Twelve references described the use of some theory or conceptual model to guide the M&E (whether in data collection, development or interpretation of results), though most were not explicitly related to gender. Commonly used models and theories included the Socio-ecological model (SEM) (n=3) [12,22,23] and Social cognitive theory (SCT) (n=2) [19,24]. Kowalczyk et al.’s review on evaluating gender-based programs found that roughly 25% (n=32) of the studies included mentioned a conceptual model or theory that guided the research; the most used was SCT [34].

Five studies developed their own guiding frameworks which included a specific gender focus such as male engagement [36], gender norms [10], empowerment [25], and agency and decision-making [9,16]. Sharma et al.’s review to identify promising practices for the monitoring and evaluation of GBV interventions highlights the valuable role of a theory of change to guide evaluations, particularly those seeking to assess effectiveness [35].

In Exner-Cortens et al.’s Photovoice study, the authors followed “feminist evaluation principles” to design their community-based participatory research of an empowerment intervention for teenage boys in Canada. They designed the photo-based evaluation method to be participatory, reflexive, and action-oriented to advance feminist goals to move away from patriarchal societal structuring and the perpetuation of misogyny. Aiming to transform the young men into active participants in the effort

to reduce gender-based disparities in health through reflecting critically on gender inequity and how to respond to it, they utilized an evaluation that centered gender [27].

#### Incorporation of gender scores into data collection and analysis

Gender scores, or the use of gender variables to quantify the impact of gender as a social construct, measure and elucidate different ways in which gender norms, roles, and relations manifest and associated impact on a particular outcome or outcomes [41]. As the way in which gender manifests is multi-faceted and context specific, a standardized set of gender variables is not relevant for all studies. Composite gender scores are also sometimes used. As such, many kinds of gender scores exist, with studies often developing their own gender scores relevant for the topic.

Twelve articles employed a gender score tool in their data collection. Often, studies would make use of an established gender score tool, most of which measured Likert scale responses. For example, the Gender Equitable Men (GEM) scale, a theoretically based measure of inequitable and equitable gender norms within sexual and intimate relationships was used in four evaluations [14,32,36,38]. This scale was originally developed by Pulerwitz et. al for their study on AIDS interventions in the United States. The GEM scale was also described in Mandal et al. systematic review on measures of women's empowerment and related gender constructs in family planning and maternal health program evaluations [37]. Similarly, Krishnan et al. used the Gender Equity Scale for Women (GESW), an adaption of the GEM, for their study on workplace gender equity in India [19]. Santhya et al. created three additive indices by combining indicators across three domains of interest (index of gender role attitudes and notions of masculinity; index of attendees rejecting men's controlling behaviors; index of attitudes rejecting violence against women and girls). These indicators were captured using items from the GEM, the WHO multicounty study on violence, the National Family Health survey, and other studies of violence conducted in India [14].

Figueroa et al. developed a gender attitudes scale for their analysis to evaluate an HIV prevention intervention in Mozambique [28] The scale was based on a series of 12 statements about gender roles assessed using Likert scale responses. Also focusing on gender attitudes, Seff et al. developed a caregiver gender-equitable attitude score to measure the effectiveness of a girl's education and safety intervention in the Democratic Republic of the Congo [12]. This score was similarly based on Likert scale responses to 10 statements related to gender roles and dynamics. In their evaluation of a teen sexual confidence intervention. Lecroy et al., employed girl efficacy and self-assertive behaviors measures developed in previous studies, also based on sets of Likert items [18].

Tura et al. used the previously validated Women's generalized self-efficacy scale (GSE) to evaluate agency created by a women's savings group in Mozambique [9]. GSE consists of constructive declarations such as "I can always manage to solve difficult problems if I try hard enough"; these are reported on a Likert scale. Continuing the theme of efficacy and agency-focused scores, Sahyoun et al. used both the Women's Empowerment in Agriculture Index (WEAI) and the Duke Social Support Index (DSSI) [15]. Evaluating a food and cooking-based employment intervention among refugee women in Palestine, they used the WEAI to assess empowerment of the women and the DSSI to understand the financial and social wellbeing of women through the tool's open-ended questions. In Mandal et al. Systematic review, authors pinpointed and described the sexual relationships power scale as a key tool to measure determinants or dimensions of women's empowerment [37].

Finally, Long et al. made use of The Menstruation Engagement, Self-efficacy, and Stress assessment (MENSES), a proven tool that aims to measure girls' experiences managing menses at school [17]. Used to analyze a menstrual health intervention among young adolescence in Mexico, this tool consists of 45 questions across three outcomes: school-engagement, stress, and self-efficacy.

#### Gender domains of analysis

Gender domains of analysis include the different ways in which gender power relations manifest as inequities. A gender analysis framework provides a structure for organizing information about gender power relations. Key domains that constitute these relations include access to resources; division of labor and activities; social norms, ideologies, beliefs, and perceptions; and rules and decision-making [2]. These can be incorporated into interview and survey questions, indicators, and

variables for analysis. Even though most studies did not use a specific gender framework to guide the M&E, many incorporated gender domains within their analysis, with some using specific tools to assess different gender domains, particularly gender norms.

For example, one study conducted a cultural consensus analysis (CCA) to assess norms among targeted cultural subgroups across the West Bank. Men and women answered questions across the dimensions of GBV, economic empowerment for women, and household and community dynamics of GBV [13]. The UZIKWASA project also captured changing norms and practice in relation to key areas of their programs, including early forced marriage, education support, and gender equitable parenting [26].

Blanchard et al. used composite empowerment indicators to explore associations between involvement with community mobilization programs, dimensions of self-reported empowerment (defined as power within, power with, and power over), and outcomes of HIV risk reduction [25]. The Sensemaker tool also explored empowerment and decision-making through a complex mixed methods methodology [30]. The authors note that this tool, while highly contextualized through co-development, can also be time and labor-intensive in low-literacy environments. It is also one of the few non-Likert scale approaches that were used.

Berti et al. modified a tool from the “Child Survival Technical support project” to assess mother’s perceptions of the father’s role in caring for their pregnant partners and in family planning [29]. Considering gender analysis domains, these measurements can be related to access to resources (i.e. access to male engagement and support), and division of labor.

#### *Approaches Which Support Gender Responsive M&E*

Other approaches which support the integration of gender into M&E include the use of sex disaggregated data and community-based/participatory methods. On their own these approaches would not to be considered gender responsive, however, they are reported here due to their important role in gender responsive M&E. M&E that only employs disaggregated data by sex or community-based/participatory methods without other dimensions of gender integration would have been excluded in search results.

##### *Sex disaggregation of data*

Sex disaggregation is important to determine whether outcomes differ for different gendered groups. There are three important points of note regarding sex disaggregation. Firstly, a precursor to sex disaggregation is including people of different sexes or genders within the intervention and/or M&E process itself. Secondly, not all programs or interventions will need to include sex disaggregated data as some will be focused on one sex or gender only (however there may still be need for engagement of other genders). And thirdly gender differences will likely be missed if gender domains of analysis, some of which that do not pertain to all genders, are not also included, irrespective of whether sex disaggregated data is used.

Six studies that included both men and women reported sex-disaggregated data [8,10,12,13,21,23]. The MENSES study collected some data from both boys and girls but did not present disaggregated data in their tables [17]. Alternatively, they used socio-economic status to disaggregate data among female respondents. The review conducted by Sharma et al. describes the importance of including participants of various genders and ages as well as disaggregating data based on age, sex, and disability status, to ensure GBV risk mitigation M&E effectively captures potential risks for vulnerable groups. They further add that data analysis should include triangulation of data from multiple sources to assess convergence and divergence of findings [35].

##### *Meaningful engagement and participation in gender responsive M&E*

Some definitions of gender responsive M&E include process-related considerations, such as how data is collected and analyzed, by whom, and/or whether a M&E process is inclusive and participatory of diverse voices, particularly women’s voices. Often such approaches are not described as being gendered unless in reference to feminist methodologies. Such approaches are important because they emphasize collaboration and partnership between investigators and participants, center lived experience as an important form of knowledge, and seek to empower, foster agency, and

amplify the voices of participants. This includes considerations for who is engaged and how they are engaged. For example, are gender norms, hierarchies, and power differentials accounted for when designing meetings and other activities? When applied in concert with other gender responsive actions, these approaches can inform how gender and gender-related elements can be integrated within M&E.

Seven studies described the use of community-based or participatory methods in the evaluation design and implementation [9,12,13,16,20,26,30]. Five of these studies collaborated with a local non-governmental or civil society organizations in study design [9,30] or data collection [12,13,20]. Two studies worked directly with community members to collect data and interpret data [16,26]. In Lees' et al., a male and a female community member trained as primary data collectors gave contextual information for the study. Further, because the study was iterative and dynamic, preliminary findings were discussed with participants in the early stages of the project [26].

The systematic review by Sharma et al., found that "building partnerships with key actors and stakeholders, including local groups and organizations" was important to ensuring effective M&E design and implementation [35]. The authors further argue that diversity factors must be considered, and M&E should be developed with relevant subgroups in mind, and with an intersectional approach that recognizes gender domains such as power and agency. Promising practices in community engagement included providing mechanisms for community feedback that are not unidirectional and facilitating community dialogues to obtain buy-in.

Sharma et al. also identified positionality of data collectors as an important domain [35]. Three studies described positionality considerations in the implementation and write-up of M&E. One manuscript described the positionality of the data collectors, including members of the authorship team (27). Burke et al. noted that to avoid potential harms, programs need to take steps to ensure that staff do not introduce their own biases, particularly those that reinforce inequitable gender norms (22). Seff et al. discussed intentionally using a separate tool for collecting sensitive information from girls, such as those on sexual behaviors and exposure to violence [12].

## Discussion

This review identified ways in which gender was integrated into M&E for RMNCAHN programs and interventions. By identifying programs and interventions which intentionally integrated gender into their M&E we were able to elicit specific ways in which gender was considered. These include using relevant theory, models, and frameworks to guide M&E, incorporating gender scores into data collection and analysis, and including gender domains of analysis (which were sometimes guided by frameworks). Additional approaches which support gender responsive M&E include disaggregating data by sex and using community-based and participatory methods, which are both good steps to inform further gender integration but need to be implemented alongside other gender integration activities. We note that there may be other ways in which gender can be integrated into M&E which are not captured here. We also recognize that there may be M&E that incorporated some of the approaches above but would not have been included within our review if specific gendered language was not used in their descriptions. Our aim was to identify ways in which gender was intentionally considered within M&E to work towards a common understanding of what gender responsive M&E is and how it can be operationalized.

Many M&E approaches included in the review were conducted on programs and interventions which intentionally integrated gender into their design and implementation. Many of the programs and interventions focused on topics such as gender-based violence, sexual health, family planning, and maternal health, and focused on the health and wellbeing of women and girls. Just because a program or intervention focuses on women or girls, however, does not make it gender responsive; this also applies to gender-based violence programs and interventions. Such topics, however, lend themselves well to gender integration, particularly those which are important for ensuring women and girls' wellbeing, autonomy, and safety. Access to family planning, for example, is vital to ensure women's reproductive choice, which in turn supports equal opportunity (42). Gender-based violence programs and interventions are important to ensure women's and sexual and gender minorities'

safety and security, which also help to support equal opportunity [43]. Given the nature of many of the articles which were included in the review, it became clear that programs which were more gender intentional in their design, more clearly conducted gender responsive M&E, however, gender integrated programs are not necessarily a precursor to gender responsive M&E.

Sex disaggregation is often described as a minimum requirement to conduct gender analysis, however, many RMNCAHN programs and interventions focus exclusively on women and girls, which means that sex disaggregated data may not be applicable. Programs or interventions which focus on one gender are sex specific, which as stated above, are not in and of themselves gender responsive. To include sex disaggregation (to be seen as gender responsive), RMNCAHN programs and interventions sometimes deliberately include men as key decision-makers or heads of households. Male engagement is an important component of gender integration when done correctly in a way that does not perpetuate harmful gender norms [44]. However, it can become problematic when men's roles as decision-makers and gatekeepers are reinforced at the expense of women's autonomy and choice. Regarding M&E, Sharma et al. give the example of interviewing heads of households, which can be problematic when only men are consulted [35].

For programs that are not gender specific, gender disaggregated data is a minimum requirement but not sufficient. Likewise, within gender specific programs, data should be disaggregated by relevant social stratifiers, such as age, race, marital status, etc. [2], however this is also not sufficient for gender integration. It is also not enough that disaggregated data is collected within the M&E process; it also needs to be analyzed in a disaggregated manner and fed into program adaptations and changes where appropriate. Within M&E, disaggregated data is often presented in tables with little to no discussion of the implications for programming, which negates having collected disaggregated data in the first place. Within gender responsive M&E, it is also important to recognize that gender norms change over time, and as such gender integration needs to be able to respond to shifting gender norms to produce meaningful results and ensure sustainability [45]. Conducting gender situational analyses, which are an assessment of the gender context and its implications for programming, will help to ensure that gender integration is relevant and context specific.

Gender responsive M&E intentionally integrates gender dimensions into the M&E process. Gender dimensions can be identified through the use of gender theories, models, scores, and frameworks. A program or intervention which has itself integrated gender into its design and implementation will facilitate gender responsive M&E. From our findings, it was clear that programs that were more gender intentional in their design, more clearly conducted gender responsive M&E. Where possible, programs should be designed to integrate gender, however, other ways that gender can be integrated include: (1) using gender theories, models, scores, and/or frameworks to guide tool development; (2) integrating gender domains (which can be guided by gender frameworks) into data collection and analysis; and (3) ensuring meaningful stakeholder engagement throughout the process, which includes gender intentionality in regards to who is engaged and how.

Our review is not without limitations. To be included in our search, references needed to include "gendered" terminology. It is likely that evaluations exist which could be considered gender responsive but do not use gendered terminology to describe their processes and outcomes. This also meant that many of the programs and interventions included in the review themselves integrated gender. Further, our search focused on M&E related to RMNCAHN only but there may be unique findings for M&E of health programs and interventions beyond RMNCAHN. Our search only included peer-review literature and did not capture grey literature and there may be other interesting ways in which gender has been integrated into M&E that has not been published in a peer review journal. We only included literature published in English and may have missed resources published in other languages.

## Conclusion

While gender intentional programs and interventions may facilitate gender responsive M&E, they are neither required nor sufficient. Gender responsive M&E is intentional in its conceptualization, participatory design, tool selection, data collection, analytic approaches, and

reporting and use. These findings will be used by the MAGE team to generate guidance on how to enhance M&E through gender responsive processes.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

## References

1. Tan KKH, Trehame GJ, Ellis SJ, Schmidt JM, Veale JF. Gender minority stress: A critical review. *J Homosex*. 2020 Aug 23;67(10):1471–89.
2. Morgan R, George A, Ssali S, Hawkins K, Molyneux S, Theobald S. How to do (or not to do)... gender analysis in health systems research. *Health Policy Plan*. 2016 Oct;31(8):1069–78.
3. UNAIDS. A tool for strengthening gender-sensitive national HIV and Sexual and Reproductive Health (SRH) monitoring and evaluation systems | UNAIDS [Internet]. 2016 [cited 2023 Feb 2]. Available from: <https://www.unaids.org/en/resources/documents/2016/tool-SRH-monitoring-eval-systems>
4. EIGE. gender-sensitive monitoring and evaluation | European Institute for Gender Equality [Internet]. 2016 [cited 2023 Feb 2]. Available from: <https://eige.europa.eu/thesaurus/terms/1217>
5. UN Women. Good practices in gender-responsive evaluations | Digital library: Publications | UN Women – Headquarters [Internet]. 2020 [cited 2023 Feb 2]. Available from: <https://www.unwomen.org/en/digital-library/publications/2020/06/good-practices-in-gender-responsive-evaluations>
6. UNDP. Gender Responsive Indicators: Gender and NDC Planning for Implementation [Internet]. 2020 [cited 2023 Feb 2]. Available from: <https://www.ndcs.undp.org/content/ndc-support-programme/en/home/impact-and-learning/library/gender-responsive-indicators-gender-and-ndc-planning-for-implementation.html>
7. Morgan R, Garrison-Desany H, Hobbs AJ, Wilson E. Strengthening effectiveness evaluations through gender integration to improve programs for women, newborn, child, and adolescent health. *Glob Health Action*. 2022 Jun 30;15(sup1):2006420.
8. Agam-Bitton R, Abu Ahmad W, Golan M. Girls-only vs. mixed-gender groups in the delivery of a universal wellness programme among adolescents: A cluster-randomized controlled trial. *PLoS ONE*. 2018 Jun 18;13(6):e0198872.
9. Tura HT, Story WT, Licoze A. Community-based savings groups, women's agency, and maternal health service utilisation: Evidence from Mozambique. *Glob Public Health*. 2020 Aug;15(8):1119–29.
10. Speizer IS, Zule WA, Carney T, Browne FA, Ndirangu J, Wechsberg WM. Changing sex risk behaviors, gender norms, and relationship dynamics among couples in Cape Town, South Africa: Efficacy of an intervention on the dyad. *Soc Sci Med*. 2018 Jul;209:95–103.
11. Shimpuku Y, Madeni FE, Horiuchi S, Kubota K, Leshabari SC. A family-oriented antenatal education program to improve birth preparedness and maternal-infant birth outcomes: A cross sectional evaluation study. *Reprod Health*. 2019 Jul 16;16(1):107.
12. Seff I, Falb K, Yu G, Landis D, Stark L. Gender-equitable caregiver attitudes and education and safety of adolescent girls in South Kivu, DRC: A secondary analysis from a randomized controlled trial. *PLoS Med*. 2021 Sep 28;18(9):e1003619.
13. Schuster RC, Brewis A, Ochandarena P, Abdelmonem A, Hosho S, Faqeeh K. Measuring community norms around women's empowerment in the West Bank: Opportunities and challenges of a novel approach using cultural consensus. *SSM Popul Health*. 2019 Dec;9:100489.
14. Santhya KG, Jejeebhoy SJ, Acharya R, Pandey N, Gogoi A, Joshi M, et al. Transforming the attitudes of young men about gender roles and the acceptability of violence against women, Bihar. *Cult Health Sex*. 2019 Dec;21(12):1409–24.
15. Sahyoun NR, Jamaluddine Z, Choufani J, Mesmar S, Reese-Masterson A, Ghattas H. A mixed-methods evaluation of community-based healthy kitchens as social enterprises for refugee women. *BMC Public Health*. 2019 Nov 29;19(1):1590.
16. Muraya KW, Jones C, Berkley JA, Molyneux S. "If it's issues to do with nutrition...I can decide...": gendered decision-making in joining community-based child nutrition interventions within rural coastal Kenya. *Health Policy Plan*. 2017 Dec 1;32(suppl\_5):v31–9.
17. Long JL, Haver J, Mendoza P, Vargas Kotasek SM. The More You Know, the Less You Stress: Menstrual Health Literacy in Schools Reduces Menstruation-Related Stress and Increases Self-Efficacy for Very Young Adolescent Girls in Mexico. *Front Glob Womens Health*. 2022 Apr 14;3:859797.
18. LeCroy CW, McCullough Cosgrove J, Cotter K, Fordney M. Go Grrrls: A Randomized Controlled Trial of a Gender-Specific Intervention to Reduce Sexual Risk Factors in Middle School Females. *Health Educ Behav*. 2018 Apr;45(2):286–94.
19. Krishnan S, Gambhir S, Luecke E, Jagannathan L. Impact of a workplace intervention on attitudes and practices related to gender equity in Bengaluru, India. *Glob Public Health*. 2016 Oct;11(9):1169–84.

20. Sharma S, Mehra D, Akhtar F, Mehra S. Evaluation of a community-based intervention for health and economic empowerment of marginalized women in India. *BMC Public Health*. 2020 Nov 23;20(1):1766.
21. Khoza N, Stadler J, MacPhail C, Chikandiwa A, Brahmabhatt H, Delany-Moretlwe S. Cash transfer interventions for sexual health: meanings and experiences of adolescent males and females in inner-city Johannesburg. *BMC Public Health*. 2018 Jan 10;18(1):120.
22. Burke HM, Packer C, González-Calvo L, Ridgeway K, Lenzi R, Green AF, et al. A longitudinal qualitative evaluation of an economic and social empowerment intervention to reduce girls' vulnerability to HIV in rural Mozambique. *Eval Program Plann*. 2019 Jul 23;77:1–10.
23. Abramsky T, Devries KM, Michau L, Nakuti J, Musuya T, Kiss L, et al. Ecological pathways to prevention: How does the SASA! community mobilisation model work to prevent physical intimate partner violence against women? *BMC Public Health*. 2016 Apr 16;16:339.
24. Dev R, Woods NF, Unger JA, Kinuthia J, Matemo D, Farid S, et al. Acceptability, feasibility and utility of a Mobile health family planning decision aid for postpartum women in Kenya. *Reprod Health*. 2019 Jul 8;16(1):97.
25. Blanchard AK, Mohan HL, Shahmanesh M, Prakash R, Isac S, Ramesh BM, et al. Community mobilization, empowerment and HIV prevention among female sex workers in south India. *BMC Public Health*. 2013 Mar 16;13:234.
26. Lees S, Marchant M, Desmond N. Addressing Intimate Partner Violence Using Gender-Transformative Approaches at a Community Level in Rural Tanzania: The UZIKWASA program. *J Interpers Violence*. 2021 Jul;36(13–14):NP7791–812.
27. Exner-Cortens D, Wright A, Van Bavel M, Sitter KC, Hurlock D, Carter R, et al. "To Be a Guy Is to Be Human": Outcomes of the WiseGuyz Program Through Photo-Based Evaluation. *Health Promot Pract*. 2021 Sep;22(5):659–69.
28. Figueroa ME, Poppe P, Carrasco M, Pinho MD, Massingue F, Tanque M, et al. Effectiveness of community dialogue in changing gender and sexual norms for HIV prevention: evaluation of the tchova tchova program in mozambique. *J Health Commun*. 2016 May;21(5):554–63.
29. Berti PR, Sohani S, Costa E da, Klaas N, Amendola L, Duron J. An adequacy evaluation of a maternal health intervention in rural Honduras: the impact of engagement of men and empowerment of women. *Rev Panam Salud Publica*. 2015 Feb;37(2):90–7.
30. Bartels SA, Michael S, Vahedi L, Collier A, Kelly J, Davison C, et al. SenseMaker® as a monitoring and evaluation tool to provide new insights on gender-based violence programs and services in Lebanon. *Eval Program Plann*. 2019 Dec;77:101715.
31. Bliznashka L, Becquey E, Ruel MT, Olney DK. Four-Year Effects of a 2-Year Nutrition and Gender Sensitive Agricultural Program on Women's Nutritional Status, Knowledge, and Empowerment in Rural Burkina Faso. *Food Nutr Bull*. 2022 Sep;43(3):364–75.
32. Ghanotakis E, Hoke T, Wilcher R, Field S, Mercer S, Bobrow EA, et al. Evaluation of a male engagement intervention to transform gender norms and improve family planning and HIV service uptake in Kabale, Uganda. *Glob Public Health*. 2017 Oct;12(10):1297–314.
33. Colombini M, Scorgie F, Stangl A, Harvey S, Ramskin L, Khoza N, et al. Exploring the feasibility and acceptability of integrating screening for gender-based violence into HIV counselling and testing for adolescent girls and young women in Tanzania and South Africa. *BMC Public Health*. 2021 Mar 3;21(1):433.
34. Kowalczyk S, Randolph S, Stokes S, Winston S. Evidence from the field: Findings on issues related to planning, implementing and evaluating gender-based programs. *Eval Program Plann*. 2015 Aug;51:35–44.
35. Sharma V, Ausubel E, Heckman C, Rastogi S, Kelly JTD. Promising practices for the monitoring and evaluation of gender-based violence risk mitigation interventions in humanitarian response: a multi-methods study. *Confl Health*. 2022 Mar 5;16(1):11.
36. Adamou BM, Iskarpatyoti BS, Agala CB, Mejia C. Exploring gaps in monitoring and evaluation of male engagement in family planning. [version 1; peer review: 2 approved]. *Gates Open Res*. 2019 Apr 8;3:1114.
37. Mandal M, Muralidharan A, Pappa S. A review of measures of women's empowerment and related gender constructs in family planning and maternal health program evaluations in low- and middle-income countries. *BMC Pregnancy Childbirth*. 2017 Nov 8;17(Suppl 2):342.
38. Pulerwitz J, Michaelis A, Verma R, Weiss E. Addressing gender dynamics and engaging men in HIV programs: Lessons learned from Horizons research. *Public Health Rep*. 2010 Apr;125(2):282–92.
39. Gender and Power: Society, the Person, and Sexual Politics - R. W. Connell [Internet]. [cited 2023 Feb 6]. Available from: <https://www.sup.org/books/title/?id=2532>
40. Masculinities by R. W. Connell - Paperback - University of California Press [Internet]. [cited 2023 Feb 6]. Available from: <https://www.ucpress.edu/book/9780520246980/masculinities>
41. Nauman AT, Behlouli H, Alexander N, Kendel F, Drewelies J, Mantantzis K, et al. Gender score development in the Berlin Aging Study II: a retrospective approach. *Biol Sex Differ*. 2021 Jan 18;12(1):15.

42. Mpoyi T. Family Planning and the Gendered Impacts of Crises on Women: An Effective Tool Across Sectors to Support Women's Empowerment and Build Resilience to Shocks [Internet]. 2021 [cited 2023 Feb 13]. Available from: <https://www.prb.org/resources/family-planning-and-the-gendered-impacts-of-crises/>
43. Noble E, Ward L, French S, Falb K. State of the evidence: a systematic review of approaches to reduce gender-based violence and support the empowerment of adolescent girls in humanitarian settings. *Trauma Violence Abuse*. 2017 Jan 1;1524838017699601.
44. Ruane-McAteer E, Gillespie K, Amin A, Aventin Á, Robinson M, Hanratty J, et al. Gender-transformative programming with men and boys to improve sexual and reproductive health and rights: a systematic review of intervention studies. *BMJ Glob Health*. 2020 Oct;5(10).
45. Heymann J, Levy JK, Bose B, Rios-Salas V, Mekonen Y, Swaminathan H, et al. Improving health with programmatic, legal, and policy approaches to reduce gender inequality and change restrictive gender norms. *Lancet*. 2019 Jun 22;393(10190):2522–34.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.