Bridging research integrity and global health epidemiology (BRIDGE): a pilot study on assessing research integrity and research fairness

Mink Albert Nicolas Sjamsoedin

Master of Science in International Health
2022

KIT (Royal Tropical Institute)
Vrije Universiteit Amsterdam (VU)
Amsterdam, The Netherlands

Bridging research integrity and global health epidemiology (BRIDGE): a pilot study on assessing research integrity and research fairness

A thesis submitted in partial fulfillment of the requirement for the degree of
Master of Science in International Health

by

Mink Albert Nicolas Sjamsoedin

Declaration:
Where other people's work has been used (from either a printed or virtual source, or any other source), this has been carefully acknowledged and referenced in accordance with academic requirements.

The thesis "Bridging research integrity and global health epidemiology (BRIDGE): a pilot study on assessing research integrity and research fairness" is my own work.

Signature:  ……

Master of Science in International Health (MIH)
Start date – 11 September 2017
KIT (Royal Tropical Institute)/Vrije Universiteit Amsterdam
Amsterdam, The Netherlands

August 2022

Organized by:

KIT (Royal Tropical Institute)
Amsterdam, The Netherlands

In cooperation with:

Vrije Universiteit Amsterdam (VU)
Amsterdam, The Netherlands

Table of content

List of tables and figures

Tables

Figures

Abbreviations

BRIDGE	"Bridging research integrity and global health epidemiology"
GHR	global health research
HIC	high-income country
ICHD	International Course in Health Development
IDI	in-depth interview
KIT	Koninklijk Instituut voor de Tropen
LMIC	low- and middle-income country
MD	Medical Doctor
MIH	Master of Science in International Health
MSc	Master of Science
PhD	Doctor of Philosophy
QRP	questionable research practices
RCR	Responsible conduct of research
RIRF	Research Integrity and Research Fairness
SD	standard deviation
SDG	Sustainable Development Goals

Bridging research integrity and global health epidemiology (BRIDGE): A pilot study on assessing Research Integrity and Research Fairness

ABSTRACT

INTRODUCTION
Research integrity has gained momentum over the last decades, which has lead to the development of guidelines to promote it. Research fairness is a movement that addresses equity in research partnerships between low- and middle-income countries and high-income countries. Recently, a first set of guidelines of research practices was published on incorporating both principles, the BRIDGE guidelines. This pilot study was a first attempt to assess the level of achievement of the BRIDGE standards and criteria and to identify possible barriers and facilitators to this achievement.

METHODS
For this pilot study, a mixed methods study design approach was used to assess current practices in global health research as experienced by alumni of the KIT Royal Tropical Institute in regards to the BRIDGE guidelines. Specifically, an explanatory study approach (QUAN → qual) was used, for which a quantitative assessment was done by an online survey, followed by a qualitative assessment by in-depth interviews.

RESULTS
Results showed lower levels of achievement of the BRIDGE criteria concerning themes on public availability of data, consideration of stakeholders and annotation of steps in research and possible sources for barriers and facilitators were identified in themes such as communication, resources, local context, incentives and ownership.

DISCUSSION
Research fairness seemed to be more difficult to achieve than research integrity. A push for open science is advocated and recommendations for a larger planned cross-sectional survey on research integrity and research fairness are made.

KEY WORDS: *global health*, r*esearch integrity*, r*esearch fairness, survey, mixed methods*

WORD COUNT: 12.979

Introduction

As a Master of Science in International Health (MIH) student at the Koninklijk Instituut voor de Tropen (KIT) - Royal Tropical Institute, one of the Advanced Modules I have taken was the Good Epidemiological Practice course. It was here where I was introduced, for the first time, to the concept of Research Fairness. I was aware of the concept of Research Integrity, as this has been a rising concern in the scientific community and also frequently discussed among Medical Doctors (MDs) in the last decade. And I found the presentation of the KIT study group on "Bridging research integrity and global health epidemiology" (BRIDGE) to be an inspiring effort to include both research integrity and research fairness into a new set of guiding principals for global health researchers to fall back on. So when a proposal was made to perform a pilot study for the BRIDGE study group of KIT, in order to assess the current practices and explore the possible barriers and facilitators to achieving a higher level of compliance with the proposed guidelines, I was very keen on contributing to their study efforts. This thesis is therefore written with the BRIDGE study group as its main audience in mind and it should be mentioned that preliminary results from this pilot study were communicated with them before hand-in of this thesis.

Background

High quality research is needed to provide quality evidence on which actors can build public policy. And in turn, to produce quality research, good research practices are needed as scientific research findings form the foundation for this evidence. But within the global scientific community, there has been a growing concern about the reproducibility of research findings (1). Published research findings may be found to be irreproducible as a result of issues such as improper study design and execution, bias and excessive statistical testing (2,3) yet perhaps even more so due to the prevalence of questionable research practices (QRP) (4). There is ample evidence of high prevalence of QRP and, to a lesser degree, even of scientific misconduct (often referred to as falsification, fabrication and plagiarism) (4–6). Responsible conduct of research (RCR) on the other hand, is seen as the ideal in scientific behaviour of researchers, which encompasses both research ethics and research integrity (7). Unlike research ethics, which focuses on moral principles and moral issues that may arise in the course of conducting research, research integrity focuses on research behaviour and the adherence to standards from the perspective of professional conduct (7). In response to the 'reproducibility crisis' (1), a call has been made to promote research integrity (7,8).

Efforts to promote research integrity include codes of conduct (9), principles to assess and reward RCR (10) and the development of guidelines for good epidemiological practice (11). However, in global health research (GHR) there is a complexity of systems at play (12) that should be considered, as the nature of its multidisciplinary and transnational components may complicate efforts to ensure good epidemiological practice (13). Specifically transnational issues and equity should be considered, as GHR often requires transnational research collaborations involving both low- and middle-income countries (LMICs) and high-income countries (HICs) (12). In this context, next to the concept of research integrity, another concept needs to be addressed being research fairness.

Research fairness is a movement that addresses equity in research partnerships (14), as a response to the fact that longstanding efforts to improve global health through research partnerships with LMICs have insufficiently lead to improved 'policy, practice, and products' (15). Collaborations between LMICs and HICs may show disparities between the benefits gained for their respective scientific staffs due to power imbalances in favour of HICs (16), resulting in issues such as inequities in professional advancement (17), lack of engagement of local researchers to shape research agendas (18), 'author parasitism' (19) or inequitable funding for research (20). This, in turn, may hinder the impact of GHR on the main beneficiaries, being local communities, as their interests may not be represented properly in research objectives and studies performed. As part of a growing awareness for the need to decolonize global health (21,22) and in order to reach the Sustainable Development Goals (SDG) (23), the importance of implementing research fairness in GHR should not be overlooked.

Problem statement and justification

In the European Union, there is a code of conduct for research integrity for researchers to refer to (9). However, until recently there have not been any guidelines which take into account the complexities of transnational research partnerships, specifically addressing possible power imbalances between institutions of LMICs and HICs, and little is known about the extent to which global health researchers abide to the principles of both research integrity and research fairness.

Recently, a set of guidelines and principles of research practices were published for good epidemiological practice to apply in GHR, the 'Bridging research integrity and global health epidemiology' (BRIDGE) guidelines (24). These guidelines were developed by a literature review combined with expert opinions by an e-Delphi method with the aspiration to bridge the gap between research integrity and research fairness in GHR (24) and offer a tool to assess research practices in this regard. However, as these guidelines are fairly new, there has not yet been a study performed on the current practices in relation to these guidelines. They have not been tested in actual settings and a means of assessment is yet to be developed in order to assess to what extent these guidelines are currently being complied with, and viewed by researchers involved in transnational and inter institutional research collaborations. A pilot study would be a logical first step, in order to test the manner in which guideline compliance can be properly measured and analysed. Next to testing the study design and research instruments, preliminary results of a pilot study may also provide some insights on what expected results may be and what aspects of the BRIDGE guidelines may need to be emphasised on in future studies regarding research integrity and research fairness.

Objectives

The aim of this research was to perform a pilot study, in preparation of a larger cross-sectional study, which will be performed by the BRIDGE study group on Research Integrity and Research Fairness (RIRF). This pilot was therefore performed, to assess the current practices in GHR in relation to the BRIDGE guidelines and to identify barriers and facilitators in achieving the BRIDGE standards and criteria. As this will also be the main focus of the larger cross-sectional RIRF study, an additional objective of this pilot study was to develop and evaluate an assessment tool, to learn from and to take these lessons into account when developing the research instruments for the RIRF study. These objectives were broken down into two main research objectives. The specified research objectives are stated below.

Research objectives

1. To asses the extent to which the principles of research integrity and research fairness, as described in the BRIDGE guidelines, are achieved in the experience of KIT alumni involved in GHR, and to identify the barriers and facilitators for these researchers in achieving research integrity and fairness in global health research.

    a. To quantify the achievement of research integrity and research fairness principles in GHR in the experience of KIT alumni.

    b. To identify possible barriers and facilitators in the achievement of the BRIDGE standards and criteria as experienced by KIT alumni.

    c. To make recommendations to foster research integrity and research fairness in GHR.

2. The development, implementation and evaluation of an assessment tool and research instruments, from which lessons may be learned and recommendations can be made for future research endeavours.

    a. To develop, implement and evaluate an assessment tool for the assessment of current practices in regards to the BRIDGE guidelines.

    b. To make recommendations for the planned RIRF study, based on the lessons learned from the development, implementation and evaluation of the assessment tool and research instruments.

Methods

Study design

For this pilot study, a Mixed Methods approach was used combining quantitative and qualitative study components. The goal of the quantitative component of the study was to measure levels of achievement of the criteria, in order to identify areas within the framework of the BRIDGE guidelines study participants experienced as being less complied with in current research practices. The qualitative component was used to further examine and explain why study participants viewed these identified areas as such. These components were performed sequentially, in the form of an explanatory design (QUAN → qual). The rationale for this was to first measure the level of compliance with the BRIDGE guidelines and, after initial analysis of these quantitative results, build on these findings to then further investigate qualitatively the reasons for variances found in compliance with the BRIDGE guidelines, connecting the quantitative and qualitative components (25).
The first (quantitative) phase of the study was done by an online cross-sectional survey, based on a bespoke BRIDGE guidelines questionnaire. The framework for and development of this questionnaire is discussed in further detail below. The second (qualitative) phase of the study was done by in-depth interviews (IDIs).

Framework

The BRIDGE guidelines were developed "for good epidemiological practice in global health targeted at stakeholders involved in the commissioning, conduct, appraisal and publication of global health research" (24). These guidelines consist of 42 criteria, divided over 6 standards, which were also identified as study phases that could potentially put strain on uniting research integrity and research fairness in global health epidemiology practices. The six study phases identified are of the study preparation, protocol development, data collection, data management, data analysis and dissemination and communication. Each standard has a number of criteria, which may be used as a checklist in the assessment of achieving research integrity and research fairness. For an overview of the full list of standards and criteria see Appendix 1. The use of this checklist is to aid researchers and stakeholders in aligning their research efforts when developing study designs and to raise awareness of possible barriers that may arise from transnational research partnerships.

Study setting and sampling strategies

For this study, a cohort was created of KIT alumni, who had received the MIH or Public Health – International Course in Health Development (ICHD) degree in the years 2016-2020, and had gained experience in GHR after graduation from KIT. It should be

acknowledged that the sample used for the survey would not be a representation of all researchers in GHR, but was to act as a pilot study population. Because of the sequential nature of the used Mixed Methods approach two separate sampling strategies were applied for the sampling of the study participants for respectively the quantitative and qualitative phase of the study. These were also performed sequentially as study participants for the second phase were selected based on outcomes from the first phase.

Survey sampling strategy

KIT alumni were approached by email through the KIT Course Administration office. Therefore, the number of invitees to participate in the study was dependent on the total number of KIT alumni that were registered with the KIT Course Administration office. As it was unknown which KIT alumni had gained experience in GHR before invitations were sent and to correct for any possible errors in the mailing list of the KIT Course Administration office, inclusion criteria questions were put in place at the start of the survey to exclude participants who did not meet the criteria. To ensure a high response rate, data collection efforts were taken by means of appropriate advocacy in the invitation email towards potential study participants and frequent reminders.

IDI sampling strategy

At the end of the survey, participants were asked to fill in their email address if they agreed to be contacted for follow-up of the survey by means of an interview. This was optional and not mandatory to complete the survey. From this group, a purposeful selection was made to invite to participate in an additional IDI. Specifically, the maximum variation sampling strategy (26) was used for selecting study participants from the first phase, allowing identification of a heterogenic group of study despite a relatively small sample size.

For the maximum variation sample, several selection criteria were used. The main selection criterion was the mean level of achievement of the BRIDGE standards and criteria, dividing study participants of the first phase of the study in three levels (low to high achievement). Additional selection criteria used were categorized by gender, years of experience (ranging from < 1 year to > 5 years) in GHR and the context of international collaboration. A first selection of study participants was made, based on an even distribution of these criteria. Based on the replies of participants not willing or unable to participate in the IDIs, this selection was expanded with a second selection of additional study participants, whilst aspiring to adhere to the even distribution of selection criteria. Eventually, a third selection of additional study participants led to all study participants who agreed to follow-up during the survey to be approached for IDIs. For a more detailed description of the selection process see "*In-depth interviews*" below. Another selection criterion that was considered was the level of education, making a distinction between KIT alumni who had reached the level of Doctor of Philosophy (PhD) since graduation and those who had remained Master of Science (MSc). However after initial analysis of the results of the survey, only one study participant had reached the level of PhD who also did not agree to being contacted for follow-up by means of an IDI. Therefore, this selection criterion was dropped.

Data collection

Survey

The assessment tool devised for the quantitative component of this study consisted of an online survey for which the BRIDGE guidelines checklist as shown in Appendix 1 was used as the framework.

Before this assessment tool was devised, a literature review was performed to identify pre-existing assessment tools for research integrity and research fairness. This literature review was performed using PubMed, Embase and Google Scholar. Main search terms used included but were not limited to "Research integrity" or "Research fairness" and "Assessment". MeSH terms were used appropriately for these main search terms, combining them with additional search terms, which included "tool", "compliance", "survey" and "guidelines". No restrictions were used and all search terms were provided of wide ranging synonyms. Snowballing was used when appropriate to find additional literature. Because these initial searches yielded limited results an additional search was performed in Pubmed, with very wide ranges, with the limitation "only reviews". 446 reviews were screened in Pubmed, however this also did not produce more articles. During the literature review for the tool development, no pre-existing assessment tool combining both research integrity and research fairness or only research fairness could be found. Most articles found on assessment of research practices were surveys focusing on only research integrity (27–31), and more specifically on misconduct and QRP (27,30,31).

The survey was developed as an online questionnaire, using SurveyMonkey. The BRIDGE criteria were rephrased as self-explanatory statements on research practices, to which study participants were asked to rate the level of achievement of these statements in regards to their most recent experience in GHR. Care was taken to stay true to the original criteria, by consulting the accompanying article by Alba et al. (32) with an extensive elaboration on the BRIDGE guidelines. Some were adapted more than others based on their clarity and a few criteria were split up into two statements, as they were deemed to extensive to create a single statement for study participants to rate easily. By doing so, a number of 47 statements were created to assess current practices in regards to the criteria. The statements were presented per standard in the online survey. Every standard, with its accompanying statements were presented on a single webpage, which had to be completed before study participants could move to the next page. See table 4 in the section of "Survey results" for the statements as presented in the questionnaire.

Achievement could be rated by using a unipolar 5-item Likert scale ranging from "not achieved" to "completely achieved". Although very limited literature was found, some example studies were found that used a five-point scale to rate perceived severity and frequency of different types of scientific misconduct (29) or to asses the level of (corporate) integrity (33).
In addition to the questionnaire on the BRIDGE guidelines, personal characteristics such as gender, level of education, years of experience in GHR and the country of their home research institutes were asked in order to later categorize for the maximum variation

sampling strategy. Participants were also asked whether or not their research involved international collaboration and if so, with countries from which continent.

The survey, once fully developed, was pilot-tested by researchers from the BRIDGE study group and independent colleagues for feedback and final adaptation before administering. The data was gathered and processed into STATA software for data pooling, cleaning and management.

In-depth interviews

The IDIs were performed after purposefully selecting study participants based on the main and additional criteria for selection. By means of open-ended questions the interviewees were asked to elaborate on why they believe some standards and specific criteria were achieved more than others, allowing for further investigation and understanding of variances found in the results of the survey. The IDIs were each separately prepared by means of a topic guide, in which the questions that were asked were both linked to the answers given by the individual interviewee as well as linked to the central themes that were identified after initial analysis of the survey results. For a more detailed description of the central themes see the "Results" section. Additional questions asked were about general ideas and thoughts of interviewees on what they might view as facilitating factors and barriers in achieving the standards and criteria in GHR. Although each IDI was separately prepared for by means of a topic guide, room was given for participants to expand on issues they deemed important and the IDIs would deviate from the topic guide ad hoc when appropriate.

These interviews were taken during online one-on-one Zoom meetings, which were audio recorded and transcribed verbatim for analysis using the Happyscribe online transcription service. Transcripts were then proofread and corrected for errors and personal identifiers were omitted. After transcription of the data into a text document, the data was processed and organized in ATLAS.ti software.

Analyses

Survey

For the analysis of the quantitative component of the study, a descriptive analysis was done for which mean levels of achievement and standard deviations (SDs) were used. First, the data was organized by inputting the data in the software program STATA after which the data was coded, cleaned and checked for missing data (34). The data regarding the achievement of the BRIDGE criteria were then given values by assigning a numeric score from 1 to 5 respectively from "not achieved" to "completely achieved". Numeric values for the participant characteristics were also assigned appropriately. After these preparations a descriptive analysis was conducted to analyse the data, applying means and standard deviations of the level of achievement per standard and per criteria, in order to identify central tendencies and variability. If participants had answered a question with the "Don't know/not applicable" option, this would be

considered as missing data and would not be taken into the calculation of means and standard deviations.

Subgroup analysis was done based on gender, years of experience in GHR and context of international collaboration. For the context of international collaboration, the countries of home research institutes were reclassified as either LMIC or HIC and the reported continents with which participants had collaborated were classified as the Global North or Global South. P-values were calculated for the subgroup analyses using the Wilcoxon rank-sum test, as a non-parametric test was appropriate due to the small sample size with no normal distribution (35). For all tests, the null hypothesis was taken as equal medians of the mean level of achievement per standard and a 95% confidence level was used. These tests were not performed per criteria, as the sheer amount of tests (141 additional tests) would create a greater probability of finding significant values ($p < 0,05$) based on Type I errors.

In-depth interviews

For the qualitative component of the study, an inductive method of analysis was done after the data was processed and organized using ATLAS.ti software, for further coding and qualitative analysis. Before coding, a preliminary exploratory analysis was conducted (34). After this initial exploration a coding process was conducted and codes were grouped per category before an inductive data analysis was performed to identify and describe trends and themes, which relate to the barriers and facilitators in achievement of the BRIDGE guidelines. In addition to identifying and describing these trends and themes, an effort was made to find disconfirming or negative cases to complement the themes and increase internal validity.

Ethical considerations

The Ethics Committee of KIT was approached with a request for exemption of full ethical review by means of a waiver (Waiver request S-170), which was obtained before embarking on this study (issued on November 1st 2021).

Neither I, nor the thesis or academic advisors had access to the names of the KIT alumni to whom an email invitation was sent by the KIT Course Administration office. Study participation was on a voluntary basis and participants could withdraw from participation at any moment and without explanation. In order to start the survey, invitees were required to give consent to obtain and use the information gathered by the survey. Study participants who were invited by email for a one-on-one online interview, received the second informed consent form as an attachment with this invitation and were asked to send a signed copy of the informed consent form beforehand.

All data collected was stored in password protected electronic data files, to which only I had access. Personal identifiers were deleted from the data files before discussing any data cleaning and coding processes with the thesis advisor. Only I was privy to the identity of the study participants who participated in the IDIs and personal identifiers were omitted from the transcripts before uploading them into ATLAS.ti.

Results

Survey

Survey respondents

Over the course of five weeks, from the 9th of November 2021 on, survey data was collected. 307 KIT alumni were invited to participate in the study of which 115 (37,5%) responded by starting the survey. Of these 115 respondents, 56 respondents were excluded from participation by answering the eligibility questions by stating they had either not received a Masters degree from KIT (n=1) or had not gained any working experience in GHR since graduation (n=55). Nine respondents gave consent but did not start the survey and did not answer the questions on eligibility. Hence, 50 respondents were included and started the survey. Of these 50 respondents four respondents opted out after inclusion, not answering any of the questions on study participant characteristics, and 10 respondents did answer the questions on study participant characteristics but did not start the questionnaire on the BRIDGE criteria statements. Ultimately, 36 respondents who were found eligible actually started the questionnaire, of which 30 completed it (figure 1).

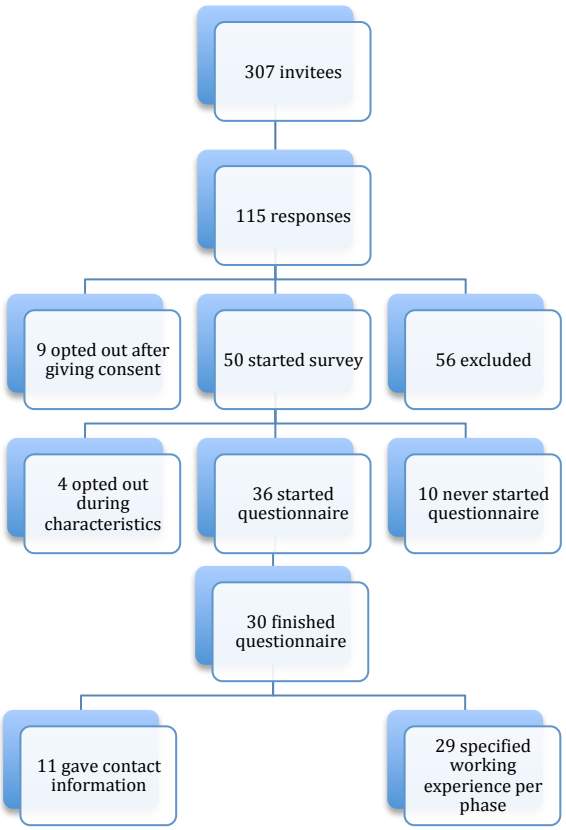

Figure 1. Flow diagram of invitation and participation process for the survey.

Participant characteristics

Of the 36 participants who started the questionnaire, 18 (50%) were men and 17 (47%) were women. The majority of study participants, both those who started and completed the questionnaire, had between 0 and 3 years of working experience in GHR (respectively 86% and 87%). The home institutions of study participants were located in and varied over 17 reported countries, of which Ghana and the Netherlands were mostly represented (n= 6), followed by Uganda (n= 4), Nigeria (n= 3), Bangladesh and Indonesia (n= 2). Two participants stated to have their home institutions in two countries (Singapore/Australia and Netherlands/USA), which were annotated as HIC's for the purpose of determining the context of international collaboration. Twenty-seven (75%) of the participants' experiences in GHR involved international collaboration and 9 (25%) of participants' experiences in GHR did not involve international collaboration. See the table (table 1) below for a full overview of study participant characteristics.

Of the 30 participants who completed the entire questionnaire, the participants had the most experience with study phases concerning the study preparation (79%), protocol development (55%) and data collection (69%). Less than half of study participants had experience with study phases concerning data management, data analyses and dissemination and communication (table 2).

Table 1: Study participant characteristics

Characteristics		Sarted Questionaire N = 36 (100%)*	Completed Questionaire N = 30 (100%)*
Gender	Man	18 (50,0%)	16 (53,3%)
	Woman	17 (47,2%)	13 (43,3%)
	Other	0	0
	Wish not to disclose	1 (2,8%)	1 (3,3%)
Degree from KIT	MIH	2 (5,6%)	2 (6,7%)
	ICHD	34 (94,4%)	28 (93,3%)
Level of education	MSc	35 (97,2%)	29 (96,7%)
	PhD	1 (2,8%)	1 (3,3%)
Years of experience in GHR	< 1 year	8 (22,2%)	8 (26,7%)
	1 - 2 years	14 (38,9%)	11 (36,7%)
	2 – 3 years	9 (25,0%)	7 (23,3%)
	3 – 4 years	1 (2,8%)	1 (3,3%)
	4 – 5 years	2 (5,6%)	2 (6,7%)
	> 5 years	2 (5,6%)	1 (3,3%)
Country of home institute	None	1 (2,8%)	1 (3,3%)
	Ghana	6 (16,7%)	5 (16,7%)
	India	1 (2,8%)	1 (3,3%)
	Nigeria	3 (8,3%)	2 (6,7%)
	Singapore/Australia	1 (2,8%)	1 (3,3%)
	Mozambique	1 (2,8%)	1 (3,3%)
	Netherlands	6 (16,7%)	5 (16,7%)
	Indonesia	2 (5,6%)	1 (3,3%)
	United Kingdom	1 (2,8%)	1 (3,3%)
	Myanmar	2 (5,6%)	2 (6,7%)
	The Gambia	1 (2,8%)	1 (3,3%)
	Uganda	4 (11,1%)	4 (13,3%)
	Kenya	1 (2,8%)	1 (3,3%)
	Yemen	1 (2,8%)	1 (3,3%)
	Angola	1 (2,8%)	1 (3,3%)
	Bangladesh	2 (5,6%)	1 (3,3%)
	Netherlands/USA	1 (2,8%)	0
	Norway	1 (2,8%)	1 (3,3%)
International collaboration	Yes	27 (75,0%)	22 (73,3%)
	No	9 (25,0%)	8 (26,7%)
Context of international collaboration**	No int. collaboration	9 (25,0%)	8 (26,7%)
	LMIC/ Global North	12 (33,3%)	9 (30,0%)
	HIC/ Global South	9 (25,0%)	7 (23,3%)
	LMIC/ Global South	5 (13,9%)	5 (16,7%)
	HIC/ Global North	1 (2,8%)	1 (3,3%)

Table 1. Characteristics of study participants as derived from survey.
LMIC/Global North = participants who have had experience with international collaboration in the context of working in a Low- and Middle Income country in collaboration with countries from the Global North.
HIC/Global South = participants who have had experience with international collaboration in the context of working in a High Income country in collaboration with countries from the Global South.
HIC/Global North = participants who have had experience with international collaboration in the context of working in a High Income country in collaboration with countries from the Global North.
LMIC/Global South = participants who have had experience with international collaboration in the context of working in a Low- and Middle Income country in collaboration with countries from the Global South.
** Due to rounding of percentages, percentages per characteristics do not always exactly add up to 100%.*
*** Ghana, India, Nigeria, Mozambique, Indonesia, Myanmar, The Gambia, Uganda, Kenya, Yemen, Angola and Bangladesh were classified as LMIC. The Netherlands, Singapore/Australia, United Kingdom, Netherlands/USA and Norway were classified as HIC. The participant, who answered the question on country of home institute with "None", also stated that their research did not involve international collaboration and therefore had no influence on the number of participants when reclassified for context of international collaboration.*

Table 2: Personal experience with research phases of study participants

Phases of research	Number of participants (%) N = 29 (100%)
1. Study preparation	23 (79%)
2. Protocol development	16 (55%)
3. Data collection	20 (69%)
4. Data management	14 (48%)
5. Data analyses	14 (48%)
6. Dissemination and communication	11 (38%)

Table 2. Phases of research study participants had personally been involved in.

Survey findings

When looking at trends in the mean level of achievement per study phase, it shows that the overall level of the achievement was reported to be higher in the Data Collection phase (Mean 4; SD 1,2) than in other phases of conducting research. The second highest mean level of achievement was found in the Study Preparation phase (Mean 3,8; SD 1,1). The study phase with the lowest mean level of achievement was that of the Dissemination and Communication phase (Mean 3,2; SD 1,5).
Overall, nearly all criteria had mean levels of achievement between 3 ("Partially achieved") and 4 ("Mostly achieved"). Twenty-eight criteria had a mean level of achievement between 3,5 and 4 and 12 criteria between 3 and 3,5.
The highest level of achievement, of 4 ("Mostly achieved") or higher, was found within 4 of the BRIDGE criteria, which were all within the standard of the Data Collection Phase. These criteria were on the use of research instruments, stating that chosen instruments were valid and reliable (after performing a review of existing instruments and their properties), were ensured to be locally adapted and culturally appropriate, for which concrete guidance for data collection was provided in a document that was available to all data collection staff and that data was collected in a respectful and safe manner and in an environment which safeguards the confidentiality of respondents.
The lowest level of achievement, below 3 ("Partially achieved"), was found in 3 of the BRIDGE criteria, which were divided between the Protocol Development phase and the Dissemination and Communication phase. These 3 criteria were all linked to the public availability of either research protocols or data. The criterion with the lowest level of achievement in the Protocol Development phase, was on the intention to make study protocols publicly available, either on a publicly accessible website or in appropriate study registers. The criteria with the lowest level of achievement in the Dissemination and Communication phase, stated that on study completion reanalyses of the data by local researchers was encouraged as much as possible and that key stakeholders and research partners were consulted to identify strategies to encourage reanalyses of the data by local researchers. The mean levels of achievement and standard deviations per criterion of the 6 standards of the BRIDGE guidelines are shown in the table and accompanying figure below (table 4, figure 2).

Table 3: Mean level of achievement per statement

Statements per ***standard***		Mean	SD	n	n NA (%)
Study Preparation Phase (Mean 3,8 ; SD 1,1)					
s1c1	Research was planned and executed in partnership with local researchers, whilst considering current professional needs and ambitions of those involved locally.	3,90	1,04	36	5 (14%)
s1c2	Key stakeholders, including representatives of affected populations and end-users of research, were identified and engaged with consideration of their needs, competences and expectations.	3,48	1,23	36	3 (8%)
s1c3*	Knowledge gaps were established by searching the literature (peer-reviewed publications and grey literature).	3,79	1,02	36	3 (8%)
s1c4*	Knowledge gaps were also established by consulting (local) experts, representatives of affected populations and end-users.	3,51	1,20	36	3 (8%)
s1c5	Research questions and objectives were developed in consultation with research partners and expected end-users.	3,55	1,15	36	3 (8%)
s1c6	Study design and research methods were selected to best fulfil the study objectives and give due consideration to multidisciplinary approaches.	3,91	0,98	36	3 (8%)
s1c7	Before embarking on primary data collection, it was assessed whether existing data could be used, fully or partly, to fulfil the research objectives.	3,97	1,17	36	5 (14%)
s1c8	It was ensured that all research partners had agreed on data ownership and publication agreements.	3,87	1,18	36	5 (14%)
s1c9	Work plans and decision-making processes were clarified and agreed on with all study partners.	3,91	1,04	36	3 (8%)
Protocol Development Phase (Mean 3,5 ; SD 1,2)					
s2c1	A detailed research protocol was prepared in consultation with all research partners.	3,81	1,19	33	2 (6%)
s2c2	A clear and comprehensive analysis section was written.	3,74	1,06	33	2 (6%)
s2c3	Studying the effect of locally relevant equity dimensions was considered.	3,29	1,04	33	2 (6%)
s2c4	When conducting multidisciplinary research, the purpose and strategies to integrate different analytical methods was described in the protocol.	3,45	1,18	33	2 (6%)
s2c5	It was strived for to make study protocols publicly available, either on a publicly accessible website or in appropriate study registers.	2,93	1,41	33	5 (15%)
s2c6	For all data collection and data use concerning human subjects, ethical approval (or a waiver) was obtained from all institutions and countries involved in the protocol.	3,86	1,22	33	4 (12%)
s2c7	When working in a setting without ethical review boards or review boards with limited epidemiological capacity, endeavours were made to strengthen local research capacity.	3,33	1,40	33	9 (27%)
s2c8	Any data sharing with third parties was explicitly stated in the protocol submitted for ethical review and in the informed consent documents.	3,83	1,28	33	4 (12%)

Data Collection Phase (Mean 4 ; SD 1,2)					
s3c1	Valid and reliable research instruments were chosen, after performing a review of existing instruments and their properties.	4,11	1,13	32	4 (13%)
s3c2	It was ensured that research instruments are locally adapted and culturally appropriate.	4,10	1,14	32	3 (9%)
s3c3	Concrete guidance for data collection was provided in a document that was available to all data collection staff.	4,00	1,15	32	1 (3%)
s3c4*	Data collection staff was selected according to technical as well as cultural criteria.	3,93	1,20	32	2 (6%)
s3c5*	The roles and responsibilities for each person involved were clarified for which adequate training and support was provided.	3,93	1,26	32	2 (6%)
s3c6	All research instruments were pilot-tested and, if possible, field-tested prior to the start of effective data collection.	3,81	1,27	32	5 (16%)
s3c7	Data was collected in a respectful and safe manner and in an environment which safeguards the confidentiality of respondents.	4,29	1,21	32	4 (13%)
s3c8	Quality assurance and control mechanisms were put in place to ensure data accuracy, completeness and coherence.	3,84	1,186	32	1 (3%)
Data Management Phase (Mean 3,7 ; SD 1,3)					
s4c1	Data management procedures were put in place before effective start of data collection and concrete guidance was provided in a document available to all data management staff.	3,76	1,21	31	2 (6%)
s4c2	A data entry application was created and pre-tested prior to effective start of data collection.	3,67	1,44	31	4 (13%)
s4c3	All variables were described in a codebook.	3,41	1,27	31	2 (6%)
s4c4	Quality assurance and control mechanisms were put in place to ensure data accuracy, completeness and coherence.	3,62	1,24	31	2 (6%)
s4c5	All data cleaning and processing steps were annotated and reproducibility was strived for by means of stored programming code.	3,41	1,39	31	4 (13%)
s4c6	For each data file, levels of anonymisation and privacy protection were defined as well as corresponding access rights in line with national and international frameworks.	3,89	1,25	31	4 (13%)
s4c7*	At the beginning of the study, an electronic secured study file was prepared to store all study documentation and outputs.	3,67	1,27	31	4 (13%)
s4c8*	The electronic secured study file was (or is planned to be) archived at the end of the study.	3,86	1,21	31	3 (10%)
s4c9	Source data was retained safely, in their original form, preserving data confidentiality for as long as has been described in the protocol.	3,93	1,33	31	4 (13%)
Data Analyses Phase (Mean 3,6 ; SD 1,2)					
s5c1	Only personal identifiers that are necessary to answer the research questions were worked with.	3,88	1,30	30	6 (20%)
s5c2*	Statistical analyses were conducted in accordance with the	3,88	1,30	30	6 (20%)

	protocol.				
s5c3*	When statistical analyses did deviate from the protocol, this was annotated and a distinction was made between pre-planned and exploratory analyses.	3,45	1,18	30	8 (27%)
s5c4	All analysis steps were fully annotated and reproducibility was strived for by means of programming code.	3,24	1,14	30	9 (30%)
s5c5	In multidisciplinary studies, statistical analyses with analyses from other study disciplines were integrated in an iterative process to coherently address the research objectives.	3,36	1,26	30	8 (27%)
s5c6	Quality assurance and quality control mechanisms were put in place to ensure that data has been correctly analysed.	3,71	1,20	30	6 (20%)
Dissemination and Communication Phase (Mean 3,2 ; SD 1,5)					
s6c1	User-specific dissemination and communication plans were developed in consultation with key stakeholders, which included (amongst others) representatives of the affected populations and end-users.	3,44	1,36	30	5 (17%)
s6c2	Data was reported in a non-stigmatising, non-discriminatory, culturally sensitive and non-identifying manner.	3,64	1,60	30	5 (17%)
s6c3	Reporting guidelines were conformed to, for the given study design and methods in academic publications.	3,50	1,50	30	6 (20%)
s6c4	Quality assurance and quality control mechanisms were put in place to ensure complete, accurate, accessible and interpretable data reporting.	3,36	1,41	30	5 (17%)
s6c5	Indexed open access journals were considered for scientific publications.	3,17	1,44	30	7 (23%)
s6c6*	On study completion, reanalyses of the data by local researchers was encouraged as much as possible.	2,48	1,59	30	7 (23%)
s6c7*	On study completion, key stakeholders and research partners were consulted to identify strategies to encourage reanalyses of the data by local researchers.	2,39	1,53	30	7 (23%)

Table 3. Questionnaire with mean level of achievement reported on a 5-item Likert scale and standard deviation (SD) per statement. The number of participants who answered the questions (n) shows a dropout of 3 participants (8%) after the first standard, followed by a single dropout per standard except for last. The number of participants who answered with "Don't know/not applicable" (n NA) is depicted with a percentage derived from the accompanying number of participants who answered that question (n). See Appendix 2 for the distribution of responses per question, which were answered with the "Don't know/not applicable" option.

** Statements that were originally part of one criterion within the standard in question from the BRDIGE checklist, but split up into 2 statements.*

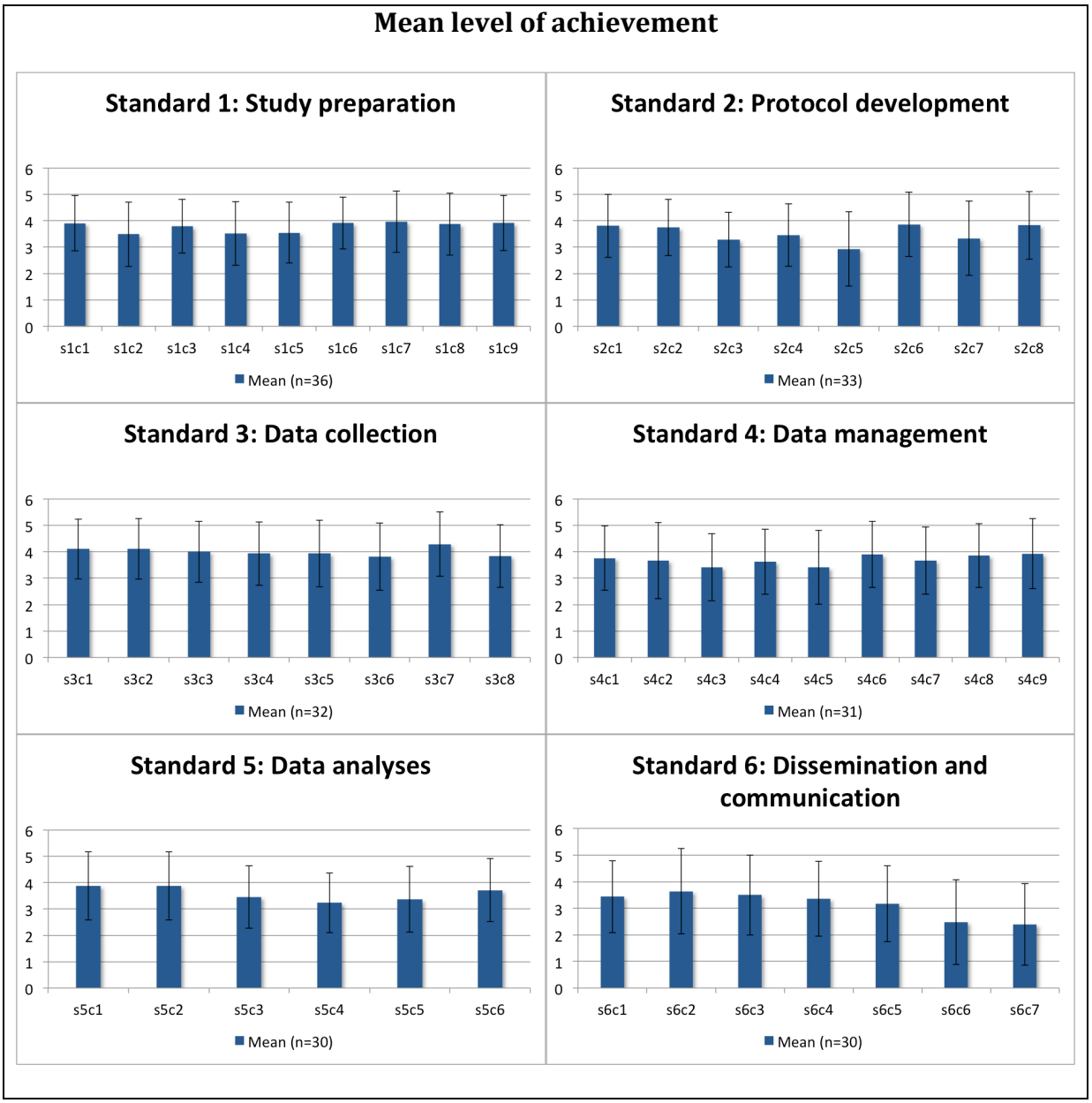

Figure 2. Mean level of achievement.
Mean level of achievement with standard deviations of all study participants.
1 = Not achieved; 2 = Slightly achieved; 3 = Partially achieved; 4 = Mostly achieved; 5= Completely achieved.
n = number of participants who answered questions per standard, which includes the number of participants who answered with "Don't know/not applicable".

Main themes identified were based on the highest and lowest levels of achievement as described above, which respectively referred to the use of valid, reliable and culturally appropriate data collection tools and to the public availability of study protocols and data for reanalyses. Other main themes that were identified, based on the 12 criteria with a mean level of achievement between 3 and 3,5, was mainly found to be on annotation (of data cleaning and processing steps, of analysis steps and deviations from the protocol and use of a programming codes or codebooks) and whether or not strategies to integrate statistical analyses in multidisciplinary studies were addressed or described in the research objectives or in the protocol. Another theme that was identified was on the consideration of stakeholders (such as the identification and

engagement of - and the development of user-specific dissemination and communication plans in consultation with - key stakeholders, representatives and end-users and whether studying the effect of locally relevant equity dimensions was considered).

Gender

When comparing the mean level of achievement between men and women it seems that women overall reported a higher level of achievement for the majority of the criteria (figure 3). Thirty-seven of the criteria (79%) were reported by women to have a higher level of achievement than the men. It is noteworthy that 4 out of the 10 criteria, in which women did not report a higher level of achievement, were on the consideration of stakeholders (s1c1, s1c2, s1c4 & s1c5). Furthermore, the differences between men and women seem to be growing larger for the Data Analyses phase and Dissemination and Communication phase. This difference was found to be statistically significant. The median of the mean level of achievement for the standard of the Dissemination and Communication phase showed a median of 3,0 for men compared to 3,9 for women (P=0,036) (table 4). No other statistically significant differences between the groups were found.

Experience in GHR

When comparing the mean level of achievement between the study participants with less or more than 3 years of experience in GHR, higher levels of achievement were reported for all standards and criteria by those with more than 3 years of experience (figure 4). This seems like a distinctive difference between the two groups, however the amount of study participants with more than 3 years of experience was considerably smaller and no statistically significant differences between the groups were found (table 4).

Context of international collaboration

A comparison of the mean level of achievement between researchers whose home institute is located in LMIC and have experience in research collaborations with the Global North (LMIC/Global N) and researchers whose home institute is located in HIC and have experience in research collaborations with the Global South (HIC/Global S) was also made (figure 5). Striking trends were not identified when comparing these two groups. The only apparent trend that could be identified would be for the Data Analyses phase, in which it seems the researchers from HIC have reported higher levels of achievement for the criteria than the researchers from LMIC. No statistically significant differences between the groups were found (table 4).

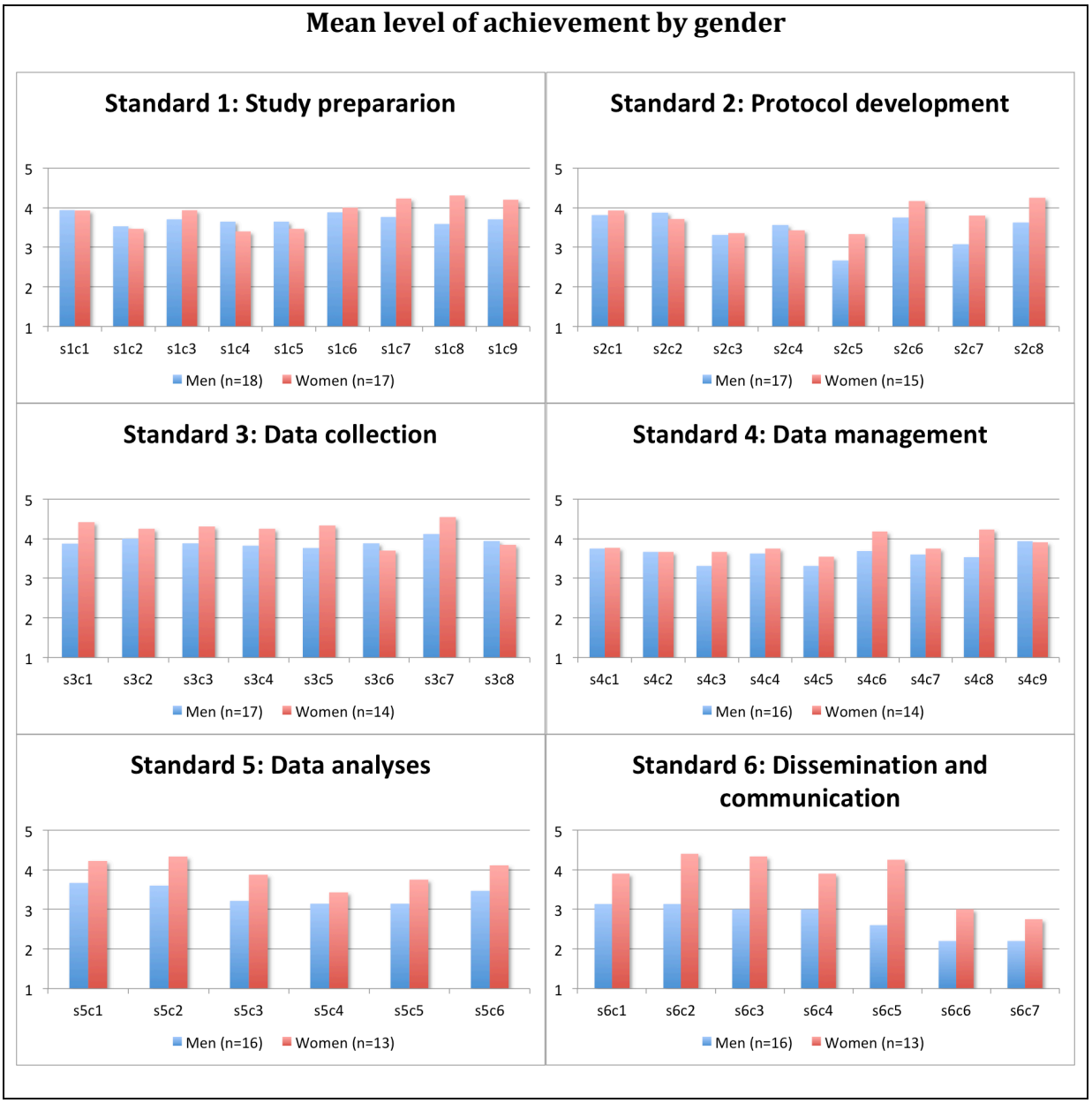

Figure 3. Mean level of achievement by gender.
Mean level of achievement, compared by men and women. Undisclosed was neglected for this comparison as n=1 for this group.
1 = Not achieved; 2 = Slightly achieved; 3 = Partially achieved; 4 = Mostly achieved; 5= Completely achieved.
n = number of participants who answered questions per standard, which includes the number of participants who answered with "Don't know/not applicable".

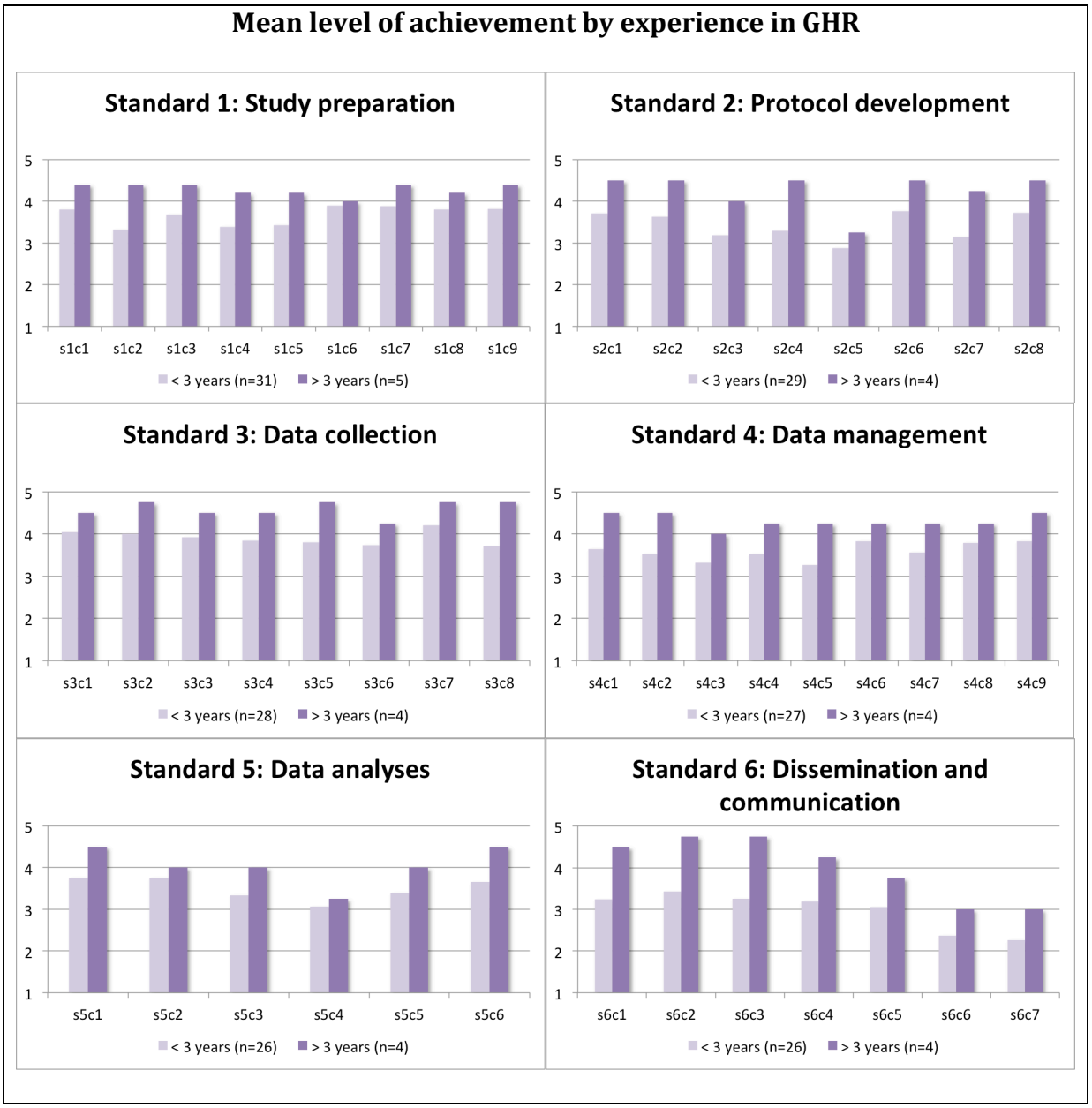

Figure 4. Mean level of achievement by experience in GHR.
Mean level of achievement, compared by participants with < 3 years and > 3 years of working experience in GHR.
1 = Not achieved; 2 = Slightly achieved; 3 = Partially achieved; 4 = Mostly achieved; 5= Completely achieved.
n = number of participants who answered questions per standard, which includes the number of participants who answered with "Don't know/not applicable".

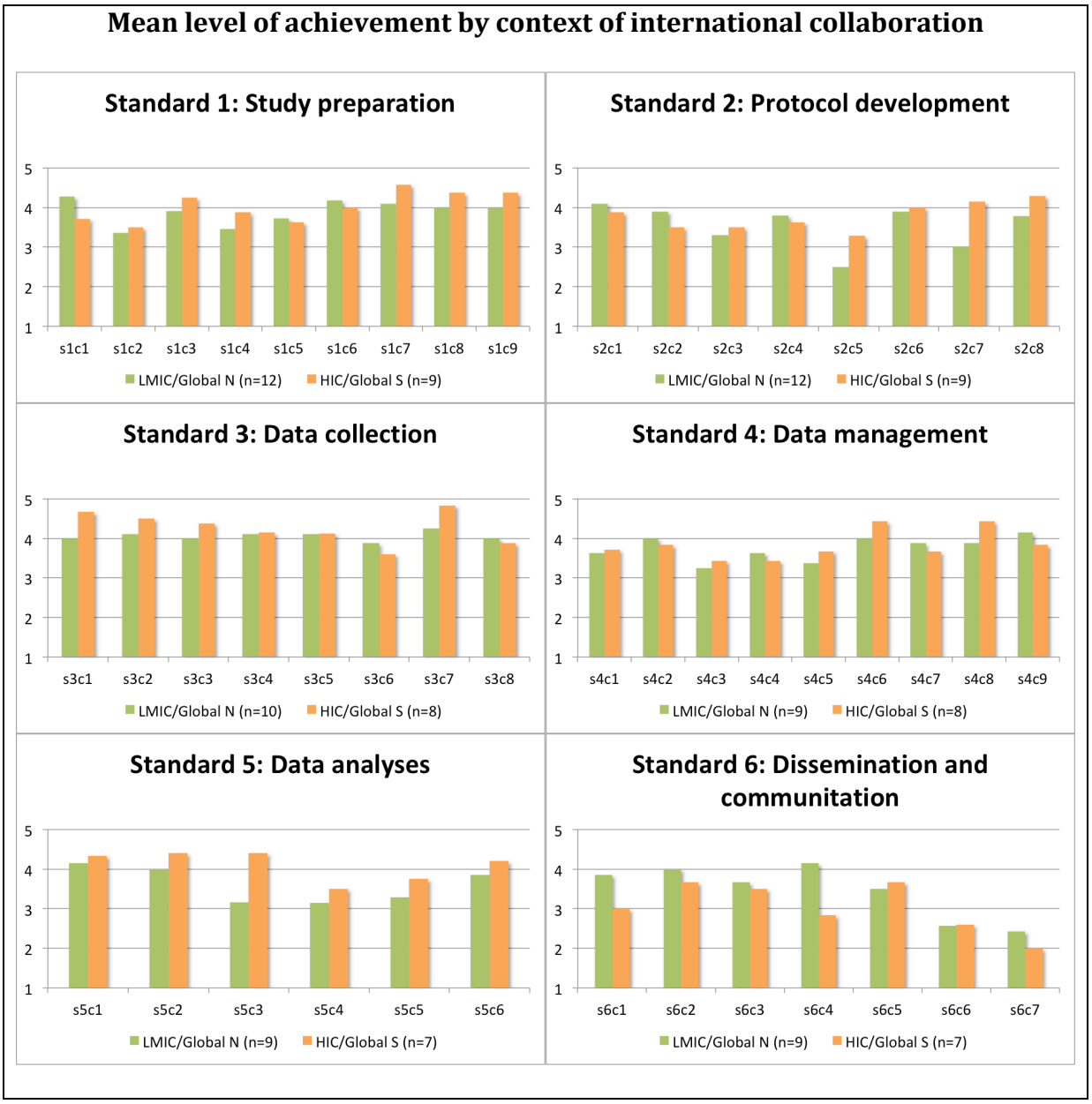

Figure 5. Mean level of achievement by context of international collaboration.
Mean achievement per subgroup of participants who have had experience with international collaboration in the context of working in a Low- and Middle Income country in collaboration with countries from the Global North (LMIC/Global N) and participants who have had experience with international collaboration in the context of working in a High Income country in collaboration with countries from the Global South (HIC/Global S).
1 = Not achieved; 2 = Slightly achieved; 3 = Partially achieved; 4 = Mostly achieved; 5= Completely achieved.
n = number of participants who answered questions per standard, which includes the number of participants who answered with "Don't know/not applicable".

Table 4: Statistical significance for subgroup analyses

Subgroup	Median (p25-p75)		P-value
Gender	**Men**	**Women**	
Standard 1	3.67 (3.33 - 4.44)	3.89 (3.56 - 4.22)	0.6344
Standard 2	3.50 (3.13 - 4.13)	3.87 (3.25 - 4.43)	0.4738
Standard 3	4.00 (3.75 - 4.88)	4.38 (4.13 - 4.75)	0.4479
Standard 4	3.94 (3.00 - 4.53)	4.11 (3.78 - 4.33)	0.6415
Standard 5	3.50 (3.00 - 4.33)	4.08 (3.83 - 4.60)	0.1166
Standard 6	3.00 (1.00 - 3.86)	3.93 (3.57 - 4.20)	0.0364
Experience	**< 3 years**	**> 3 years**	
Standard 1	3.78 (3.33 - 4.22)	4.67 (4.11 - 4.78)	0.0902
Standard 2	3.67 (3.06 - 4.17)	4.38 (3.69 - 4.81)	0.1640
Standard 3	4.25 (3.38 - 4.75)	4.69 (4.25 - 4.94)	0.2511
Standard 4	3.93 (3.00 - 4.44)	4.17 (3.94 - 4.67)	0.2948
Standard 5	4.00 (3.00 - 4.33)	4.00 (3.50 - 4.33)	0.7834
Standard 6	3.43 (2.00 - 4.00)	3.93 (3.50 - 4.50)	0.1972
Context	**LMIC/Global N**	**HIC/Global S**	
Standard 1	3.89 (3.56 - 4.22)	4.17 (3.33 - 4.71)	0.5309
Standard 2	3.71 (3.00 - 4.33)	3.88 (3.49 - 4.55)	0.4992
Standard 3	4.25 (4.00 - 4.75)	4.54 (3.69 - 4.94)	0.6533
Standard 4	4.06 (3.50 - 4.33)	4.22 (2.50 - 4.72)	0.7770
Standard 5	4.00 (3.50 - 4.33)	4.30 (3.83 - 5.00)	0.4674
Standard 6	3.86 (3.14 - 4.20)	3.64 (1.71 - 4.20)	0.8112

Table 4. P-values were calculated per standard for subgroup analyses using the Wilcoxon rank-sum test. For all tests, the null hypothesis was taken as equal medians of the mean level of achievement per standard. A 95% confidence level was used.

In-depth interviews

IDI participant selection

Of the 30 study participants who completed the survey, 11 (37%) provided their email addresses at the end. Of these 11 IDI candidates, six were initially invited for the online interview based on the maximum variation sampling strategy (table 5). Of these six, three agreed to participate in the IDI, two declined follow-up and one email address was invalid. After this first round of invites, a second selection was made, including three more study participants who would (if agreed to the IDI) satisfy the evenly distributed selection criteria of the whole group of interviewees. After no replies were received after the second reminders, final reminders were sent and the remaining two participants were also invited for IDIs. This process took place over the course of eight weeks, from the 7th of April 2022 on, in which four IDIs were eventually taken (figure 6).

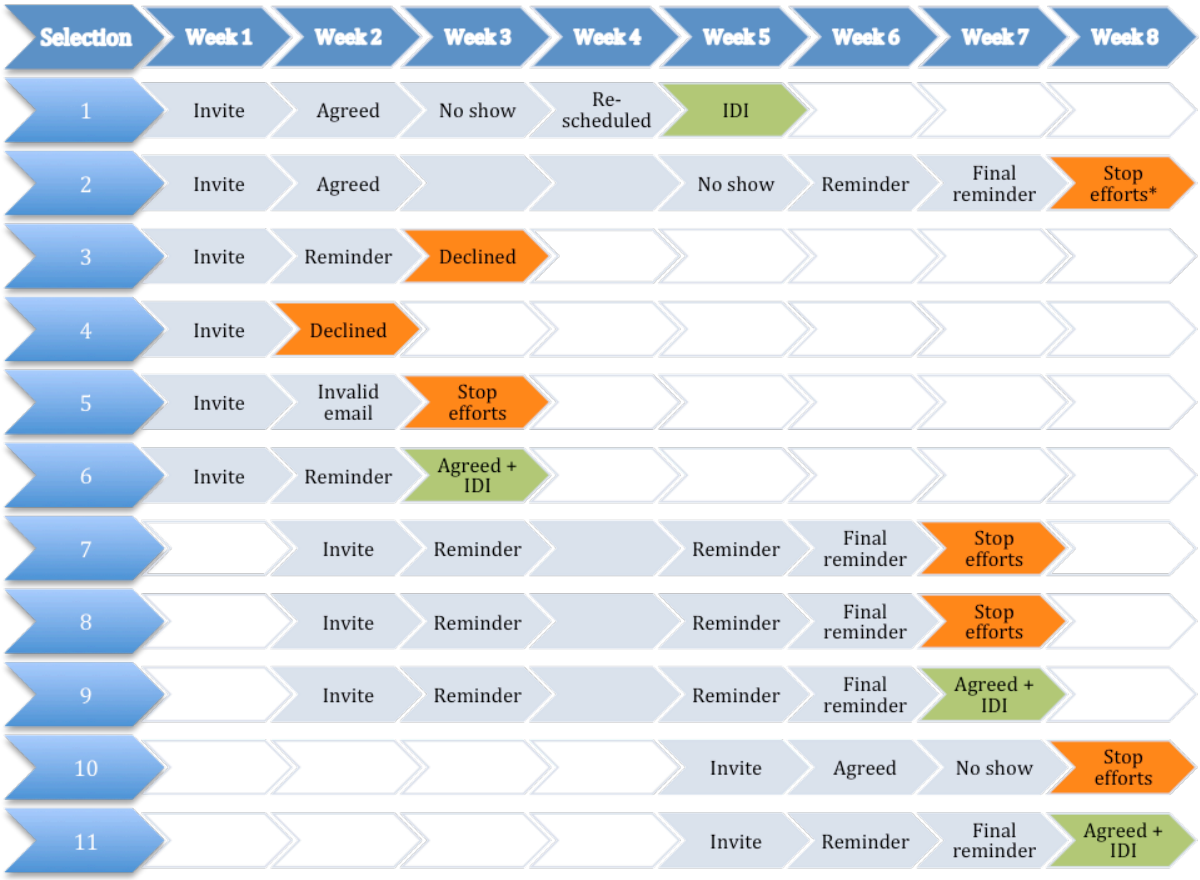

Figure 6. Timeline for the IDI invitation process.
The IDI candidates were invited for IDIs and, if agreed on, planned for. After a non-response, participants would receive a second invitation (reminder) after one week. Second reminders were sent 10 days after the first, with a third reminder after one more week if no response was given still. Participants, who agreed to and planned an IDI but did not appear for the IDI (no show), were sent a new reminder. Efforts to plan an IDI were stopped if contact information appeared invalid, after 3 reminders were sent without response or if re-establishing contact was unsuccessful after a no show.
** Contact was re-established, but rescheduling of a new meeting was unsuccessful.*

Table 5: IDI candidate selection criteria

Category/IDI candidate	**1**	2*	3*	4*	5*	**6**	7*	8*	**9**	10*	**11**	**Sum**
Mean achievement												
Highest	**x**				x		x			-**		**1**
Intermediate				x		**x**		x		-**	**x**	**2**
Lowest		x	x						**x**	-**		**1**
Gender												
Man	**x**		x		x		x	x			**x**	**2**
Woman				x		**x**			**x**	x		**2**
Undisclosed		x										**0**
Experience												
< 1 year		x		x			x	x		x		**0**
1 – 2 years	**x**				x				**x**		**x**	**3**
2 – 3 years			x			**x**						**1**
International Collaboration												
None	**x**											**1**
LMIC/Global N			x				x			x	**x**	**1**
LMIC/Global S						**x**			**x**			**2**
HIC/Global N				x								**0**
HIC/Global S		x			x			x				**0**

Table 5. Distribution of selection criteria of the 11 IDI candidates who agreed to participate in IDI during survey, based on mean achievement and characteristics for selection of interviews. 1 – 6 were part of the first selection, 7 - 9 were part of the second selection and 10 & 11 were part of the third selection. The sum of 4 IDI candidates was interviewed, showing the actual final distribution of participant characteristics.
** IDI candidates who were invited but not willing or unable to schedule an IDI.*
*** IDI candidate had answered all questions of the survey with "Don't know/not applicable".*

IDI findings

Four of the IDI candidates were interviewed for qualitative data collection and analysis. From these interviews six main themes were identified that could be considered as a main source of barriers and/or facilitators in achieving the criteria of the BRIDGE guidelines. These main themes will be discussed in the subheadings below.

During the IDIs it was observed that interviewees did not always have a recollection of the answers they had given during the survey and that they would sometimes want to amend their given responses after more explanation on the BRIDGE criteria. Interviewees would often give more general answers to the questions asked during the interview and when specific criteria were discussed it sometimes seemed to be unclear to interviewees what was exactly meant by some of the statements. This could sometimes be due to the choice of words, which participants were not familiar with, or due to the complexity of the statements. It should be noted that for all four interviewees, English was not their native language.

1) Communication

All interviewees stated that one of the main barriers in achieving the criteria were founded in a lack of communication in preparation of conducting research. When asked about the involvement of key stakeholders, in both the study preparation phase and protocol development phase, some of the interviewees stated that equal collaboration was hindered because of this lack of communication.

"When I went to the village office, to the leader, there were a lot of arguments to discuss, because I didn't receive the first briefing, of course. It's not equal because the village leader, he doesn't have health education; I'm not sure. Also on the field, when I met the village midwife—there's a village midwife—she didn't know about the research for example. We have to make sure that the communication between all the stakeholders involved should be provided." (IDI candidate 9)

Besides the lack of communication between stakeholders, all interviewees also stated that the lack of briefing and formal training for the data collection staff was a barrier in achieving the BRIDGE criteria. The lack of knowledge on the subject matter and also the lack of knowledge in data collection were main themes that were identified.

"Then what I observe during the research, like when we were talking about the table, like HIV or AIDS because I found several cases in the village, the community health workers, they don't want to visit the mother. I think the community health workers and the village officers, they should be given about the correct knowledge about HIV and aids. It's like a taboo. Because they were afraid to be infected. Something like that. That's happened." (IDI candidate 9)

These barriers were not deemed unsolvable however and with additional efforts put into training and close supervision two of the interviewees stated that they had been able to overturn this.

"Actually, there is no formal training for them for the questionnaires, primary data collections, something like that. I didn't give them a formal training like a two-day or three-day or something like that. Actually, I make a coaching, a closed guiding on them, like a on-job training. I show them how to use the SPSS software. I explained about the guidelines, the research tools, and the contents and the nature of the information that we are asking from the various states and regions. I make them like a close supervisor. I work with them, they work with

me together, and I coach them. It's a very active thing of make the process very successful, I think." (IDI candidate 1)

Three interviewees also mentioned that creating an atmosphere for frequent and open communication in turn worked as a facilitator in achieving the BRIDGE criteria, that lead to better quality assurance and control mechanisms in performing research. Not only did this atmosphere have a beneficial effect on achieving the criteria on quality assurance and control mechanisms, it also seems to be linked to the understanding of local context (see subheading "Local context" below).

"I discussed intensely with my international consultant. Actually, she's like a big sister for me. We are a little family here, and we stay friendly, and we discuss very openly on the issue, like how to maintain the optimal quality of the research, the tools, and the data, and the process. I think we discussed 1 to 2 weeks before the whole process started." (IDI candidate 1)

Interviewees also identified proper communication with local communities and stakeholders as a requisite for achievement of the criteria, as it influences the acceptability of the process by study participants. One of these interviewees also made a mention on language barriers as being a hindrance to proper communication. These were also linked to the understanding of local context and will be described further below (see subheading "Local context").

"we need the community participations and all stakeholders including approach […] so that we can establish the research successfully at the implementation to be successful. Sometimes that's the community. The community has a lot of different types of people. Some people are low educated, some people are higher educated. Sometimes, the balancing their level of understanding and their level of knowledge are difficult. […] At the time, they mentioned that their report, that they have some challenges to explain because of the language barrier. At the very first time, very first man of the surveys or questionnaire appearing at that research, they faced the difficulty are to explain and do discuss with the local people at the time. After that, they realized that they needed to translate this questionnaire to the layman term, and to easily familiarize and the user-friendly to the local context, and so that they hire the translator. " (IDI candidate 6)

Some participants also mentioned COVID-19 as a hindering factor in communication, as meeting with stakeholders and collaboration with research partners was complicated. This and other COVID-19 related barriers and facilitators will be further described in the subheading "COVID".

2) Resources

A main source of barriers in achieving the BRIDGE criteria was found in a lack of resources regarding time, budget and data collection staff. These were issues that would interfere with maintaining the quality of research endeavours.

"All the research and majority of the research are very expensive to maintain the gold high standard. How can I explain like that? If we want to produce, if we want to implement a research maintaining a gold standard, we have to spend a lot of money. But when I started my journey as a researcher here, it is started at just midpoint of the COVID. Other times, you know that the funding, we are depending on funding from developed countries. It shrink, the funding shrink a lot. So every researcher, every... As I said, this very expensive, so we have to maintain quality with limited resources." (IDI candidate 11)

And it was specifically mentioned that time and monetary restrictions lead to prioritization, which would not always leave room for proper assessment of existing

data, establishment of knowledge gaps by literature reviews or considering local equity dimensions.

"I think when we get the grant, the time period that we get from the donor, it was very short. For this reason, we had to prioritize. We have to give priority with area we should focus fast, then we'll move next. Maybe this is one of the reason we do not focus on that domain." (IDI candidate 11)

The same was mentioned when interviewees were asked about protocol adherence and deviations from pre-planned statistical analyses.

"It is important to have complete analysis plan during writing protocol, but what happened in […] third-world country, most of the time we submit protocol with the very general analysis plan, because many other reasons like time constraints, resource limitation. At the same time, every team don't have good data analyst. So after completing the implementation, we'll have all data. Then maybe we'll hire a consultant for one or two months to done the analysis. But before getting the data, it does not happen. It is also maybe one of the reasons." (IDI candidate 11)

The monetary resources were often mentioned to be the reason for shortages in data collection staff and were also linked to a lack of training for this staff (see "Communication"). And in consideration of not only the data collection staff but also researchers themselves, an overburden of the workload was mentioned by all interviewees when asked about maintaining quality assurance and control mechanisms, pilot-testing reliability and validity of research instruments or establishing knowledge gaps by searching literature or consulting local stakeholders. This overburden was mainly due to the many other tasks the data collection staff had, regarding their work as health care providers.

"Another thing is they are overburden of the workload because our busy house staff they are doing very many tasks. This nationwide program asked them to do so and that nationwide program asked them to do so. They are like end products of every national program. So they are doing some many times and they have to fill out many sheets from different programs. It is like an overburden of the workload. That makes incompleteness of the information. When we looked at the sheets and the data and the books, there are so many gaps they didn't fill out because some say that they have no time to fill out the forms and something like that." (IDI candidate 1)

Two of the interviewees also mentioned that budgetary issues, which lead to shortage of data collection staff, were linked to a lack of transparency as this might have *"effects to the finance" (IDI candidate 9)* (see subheadings "Local context", "Incentive" and "Ownership").

"The organizations... Actually, it is research organization, but they are actually profitable organization. They are actually business organization. If you give ten dollars for data collections, they will try to achieve as much as they want to interest. They want to save money as much as they can." (IDI candidate 11)

3) Local context

Local context was identified as one of the main sources of barriers in achieving the BRDIGE criteria. It was stated by one interviewee that a lacking of cultural prowess to data collection and data management was experienced to be a challenging factor to maintain the quality of research. Other interviewees also mentioned that local context may hinder the ideal way of data collection, as local customs would not always allow for researchers to adhere to their own protocol. A lack of knowledge as described above

(see “Communication”) could also be linked to the local context of a country, as well as issues with data collection staff (see “Communication” and “Resources”).

“And also, maintaining the high quality, it also depends on the participants sometimes. The participants and also the local leaders, local context. Sometimes if it is required to collect data or sample, after maybe, at the evening when there is no sunlight, maybe it is gold standard to collect the data after sunset, but you cannot sometimes because the society will not permit it. You have to compromise with the quality. But at that time what we will do, we'll just collect the data before the sunset. Maybe sometimes, we are able to manage the society, sometimes we fail, but we have to collect data. We have to compromise little amount of quality, but try our best.” (IDI candidate 11)

It was also noted that even when research is performed properly, it does not always give the answers needed to address the issues within the local context. This can be because of difficulties in translating research results into policy (and the extrapolation of the results to different contexts), but also has to do with the acceptability of both research endeavours and research findings.

“One thing. Gold standard, sometimes we're fixed by the peoples who are from the well-structured areas. A standard is fixed by the peoples, who are in a well-structured areas.[…] For example, the majority standard themes, it is developed by developed countries, the researchers from developed countries. It's difficult for them to see the actual scenario in the developing countries or low-developed countries. Also we do research for peoples. You produce very good quality things, but people do not accept, because they don't have enough knowledge to accept it. So what we should do? We should do the best quality for them, not best quality in a scientific quality. Not in a scientific quality. But if it's not accepted by them, then it means that... For me it is not population science, because population should accept it. Otherwise they will not developed, they will not improved. So if anything accepted by peoples, it is the gold standard, should be.” (IDI candidate 11)

This acceptability was also mentioned to be influenced by the amount of available local resources.

“Actually, they are not trying to use the products of the research for other programs, I think, because the planning and data collection are aimed for that province. And that province and other province, their local contexts are, I think, not very similar, and they are reluctant to put other additional resources to way forward. That kind of thing is very common in our bureaucratic system.” (IDI candidate 1)

However, interviewees also mentioned that these issues could be overcome by extensive communication on the subject with the various stakeholders. This links to the findings on frequent and open communication as described above (see “Communication”) and also to the benefits that local stakeholders may receive in the process (see subheading “Incentive”).

“I had the feedback. After the dissemination of the report, there are a lot of discussion and meeting at the local level and also at the central level. We review and revise the system, the whole system, the surveillance system. Then we make the modification points, which are suitable for different local contexts. So the modification points, some of the states and region, they started using the modification points and they benefit a lot.” (IDI candidate 1)

And in turn, one interviewee also stated that the acceptability of the local communities would benefit if research instruments were catered to local culture.

“I didn't find out any difficulties. Actually, it's based on the local context. They can accept for the research.” (IDI candidate 9)

One interviewee stated that unfamiliarity with local context could also lead to a misinterpretation of data. This was seen as a barrier in making data available for secondary analyses by other researchers and was also linked to ownership of data (see "Ownership").

"It's just actually very difficult question. I also see that one scientist, renowned scientist, who publicly share a Tweet that some people ask me data, and some people email me. It is his Twitter tweets. He said that some people ask me the data for meta-analysis. But I collect it and I go to the field, my field workers work a lot of time to collect the data, and the peoples who emailed me, they want data for publications. So it is his actually opinion, and he said that the people who do reanalysis, he may not know the context. He may not know the actual protocols, or actual scenario on the time during data collection. So it is also very difficult for the people who do the reanalysis who are not involved with the history." (IDI candidate 11)

In contrast to the local context being associated with as being a source for barriers in achieving the BRIDGE criteria, familiarity with local context was considered to be a facilitator in achieving these.

"Actually, the [omitted] consultant understands very much about our organization, the bureaucratic system. She is very familiar with our organization. She understands my situation, my environment, so we can discuss very well. That is one strength for the process, I think." (IDI candidate 1)

And a better integration of researchers from LMIC and HIC was mentioned as a facilitator in research collaborations and additional efforts for this integration advocated by one of the interviewees to further familiarity with local context.

"You also agree that the situation is improved a lot. Accommodation is improved, collaboration is improved, where you are talking from very far area. So it is very difficult before 20 years ago. It was very difficult, or 40 years ago. First of all, we have to interchange the researchers, the researcher who want to work for peoples from developing countries. They also should stay sometimes with that context with that people that they want to, because sometimes it's not possible. Maybe that time, the collaboration with the researchers from low-and middle- income country is possible, but one of the best thing is to stay on that community. […] Another thing, it has also happened in KIT and all other things. Researchers from low and middle-income countries for maybe 30 years or 25 years, they can work there, developed country, and on that time, they can extend his knowledge and her knowledge between the researchers from developing countries and researchers from developed countries to develop the protocol, to develop pushing it, to analyse the findings." (IDI candidate 11)

A noteworthy sentiment that was observed during the interviews, was that three of the interviewees also advocated for more flexibility when it comes to adherence to guidelines, in consideration of the local context. They felt that guidelines and standards should be adaptable to local context, whilst keeping in mind that the expected outcomes of research endeavours should be used to benefit mostly the end-users. This was also linked to the acceptability of local communities to possible outcomes and policies.

"Of course, we should follow the guidelines, maybe from the World Health Organisation or other international bodies. But I don't think we can use the general guidelines. We should consider the local context, […] ad hoc guidelines […] should be flexible." (IDI candidate 9)

Another issue that was raised by two interviewees was a lack of transparency, which may also be considered as part of local context. However, as described above (see "Resources") this may also be linked to monetary resources as well as to the involvement of stakeholders. This will be described further below (see subheadings "Incentive" and "Ownership").

"Another challenge is the transparency. It's like a culture in my country. Because some issues like tender process, the selection of the product and the price list, […] this going to be input in the calculation and they can be replaced. Some people don't want to give the exact information. […] That is a major challenge." (IDI candidate 1)

4) Incentive

Another source for facilitators in achieving the BRIDGE criteria was found in the involvement of local stakeholders. As described above (see "Communication") close supervision was found to be a facilitator. However, this was not only due to improved communication, but also due to the involvement of the data collection staff and their perceived benefit from such supervision.

"For me, I think the major factors which make the process successful is the willingness of the stakeholders and the skills and enthusiasm of my […] medical officers, […] they are very active and they are very interested in their work. […] Also for the stakeholders from the states and regions, […] That focal persons are very active, too. […] They are very active, and they advocate their superiors, and they are actively collecting the data and the information that I wanted from the central side. Then they collect the data at any hour and then give to me. When I asked them to please recheck, please reconfirm the dataset, something like that, they are willing to do on the same spot. That kind of active thing, I think that is the essential factor that makes the process successful until now. […] I think they have many benefits from the research itself and also making research together with my team, because they learn how to plan and implement the research, how to develop the research, the research methodology, and also for the statistical software. These are, I think, the benefits they get, because at the medical officer level, they are not taught them." (IDI candidate 1)

One interviewee also stated that the research institute he worked for had a lot of incentive to maintain good quality, as this was their *"main accreditation" (IDI candidate 11)* (see "Communication). This interviewee was very confident in their capacity to build in quality assurance and control mechanisms, to use valid and reliable research instruments and to insure that research instruments were locally adapted and culturally appropriate.

"Yes, we try our best to make it reliable. We do not compromise. […] I think all the time. […] it is very important for our research institute." (IDI candidate 11)

The incentives of different types of stakeholders were also mentioned as a source of possible barriers in achieving the BRIDGE criteria regarding quality assurance in data collection efforts. These barriers were also linked to monetary issues (see "Resources") and the lack of transparency (see subheading "Ownership").

"the financial, the budgetary can be accessible or not or to be online platform. What I mean is maybe it relates to the budgetary, because what I felt is the village office, they didn't expect that the report should be published because it effects to the finance." (IDI candidate 9)

Another barrier in achieving the BRIDGE criteria, specifically about making data publicly available for reanalyses, that was associated with incentive was the amount of credit that primary data collectors enjoy from well performed research. This will be described below (see subheading "Ownership").

5) Ownership

When interviewees were asked about the dissemination and communication phase and about public availability of research findings and primary data, a barrier to achieving these criteria lied in the form of publication that was used. The form of publication would often depend on who was the originator or the main funder of research and influence the public availability of publications and encouragement of reanalyses.

"Sometimes in our organization also many papers, many reports are not publicly available. The final full report not available, but in some format, it is available. It is available in, I mentioned, in government website in some projects, and it is also available in daily newspapers because our organization also published papers, published report in daily newspapers. At the same times, we also published our findings in our own organizational websites for papers, and also social medias." (IDI candidate 11)

Deviations from pre-planned analyses in the protocol were mentioned to not always be annotated due to the format of publication. Interviewees noted that if a publication was done in a journal, mostly the positive endeavours would be mentioned. But if a publication was done to a donor, this donor would get a more comprehensive report, also because it was in their interest to be able to justify how they had spent their budgets.

"Yes, as my experience, I think they present a step-by-step study and there are changes, small changes or big changes. Most of the teams, they present to the donor. But for the publication, they only show that at the final results, and I think just one page or just sometimes there are one or three page. […] I think that when I read a full paper from the medical journal, they only show what they were in, and then their final results, and then for the recommendation like that. […] the donor's interest, what project period, budget, and also project extensions and the changes, which affect on the result that don't apply for the publication in medical journal" (IDI candidate 6)

Ownership of data could, in some cases, also lead to a lack of transparency. This was often linked to the involvement of stakeholders and their willingness to provide data, which in turn was often linked to monetary motives (see "Incentives" and "Resources"). It was also mentioned that it was not always clear who was the owner of data and that this could lead to a lack of transparency in available data and therefore hinder data collection efforts.

"It's just like a secret. Not everyone knows the correct budget and the correct finance for the districts. It just kept for the village leader himself." (IDI candidate 9)

One of the interviewees also specifically mentioned the credit researchers receive for their work as a barrier to achieving the BRIDGE criteria regarding the encouragement of reanalyses.

"I also feel like sometimes when anyone collect data because everyone want to get credit. So the people who reanalyse, maybe he do not give enough acknowledgement who actually collect the data, who are the main principal investigator. When you publish papers, or when you write an article in newspapers or in for conference, people will not 100%, then everybody will give credit to you that you done very good job. But it is actually data from another person. Actually, the audience will not know, the audience will not analyse a lot of things, like you do not want to see that, actually, who is the main owner or who is the main investigator who collect that data. Who present, he is the actual owner." (IDI candidate 11)

This issue on encouragement of reanalyses was also linked to the local context, in which it was stated that the owner of a primary data collection was familiar with the local

context and another without this familiarity might misinterpret the data (see "Local context"). In contrast to this view, another interviewee stated the opposite, saying that secondary analyses should be performed without the influence of the primary data collector in fear of personal bias. This interviewee advocated that secondary data analyses should in fact be encouraged, without the influence of the primary data collector.

"I think that different people have the different opinion, different thinking, different view, so I think it should be allowed to do the secondary data analysis. If some people, like our students, they want to do the secondary data analysis, the primary owner of the research, researcher, should not be included, because I think that we should aware of the bias or personal bias. […] the primary data collection, it's also important and also secondary data collection is too. […] But if the student want to use their secondary data analysis […] the primary owner of a research owner should not be included in the secondary data analysis review process, because different people has different ideas and different view. I want to know the different things and I want to collect the different peoples from different things. Personally, I don't like the people who influenced on the other people's idea." (IDI candidate 6)

Another way in which ownership was mentioned to be a possible barrier in achieving the BRIDGE criteria, had to do with agreements between research partners on data ownership and publication agreements.

"About first author and the senior author, last author, and for screening author, most of the time, in our organization and also beyond our organizations, it is fixed. But in between, sometimes there is some flexibility or not clearly mentioned before starting or before finishing the study because something is actually unclear, because sometimes it depends on amount of time and amount of contribution to give the higher rank in publication. Before completing the papers, adopting the manuscripts, it is difficult to know how much everyone contribute." (IDI candidate 11)

6) COVID

During the interviews, there was often a mention of COVID-19 as being a barrier in achieving the BRIDGE criteria. This had to do with difficulties in communication, the decrease in funding for GHR (see "Resources") and also the fact that COVID-19 was still quite novel, which hindered efforts to establish knowledge gaps as there was little confirmed information to explore by means of a literature review.

"COVID makes trouble to get sufficient collaboration with the partners, with the external consultants, and also others. […] it is very what we call very novel study. It is also one reason that there is no existing data about that domain because COVID virus is very new, so it is maybe also one of the reasons. Also time restrictions, communication gap had also important role." (IDI candidate 11)

Discussion

This chapter is guided by the research objectives. Therefore, first the study findings will be discussed, followed by a critical appraisal of the assessment tool used to perform this study. Finally, a reflection on the execution of this pilot study will be given.

Study findings

Summary

The overall reported level of achievement of the BRIDGE standards and the majority of the criteria was between "Partially achieved" and "Mostly achieved". Themes identified within the standards that were reported to have higher levels of achievement were on data collection efforts. Themes that were identified to have lower levels of achievement were on public availability, consideration of stakeholders and annotation of steps within different phases of research. Overall, women reported higher levels of achievement than men and study participants with more than 3 years of working experience in GHR than those with less than 3 years of experience. During the IDIs, several themes were identified that could be possible sources of barriers and facilitators which could influence the level of achievement of the BRIDGE standards and criteria. These themes, which included communication, resources, local context, incentives and ownership, seem to all be connected to one another and can be considered to have both possible negative and positive effects on achieving the BRIDGE standards and criteria.

Interpretation

The two standards that had the highest mean level of achievement were on the Study Preparation phase and the Data Collection phase. And within the Data Collection phase, the criteria reported to have the highest levels of achievement were on the research instruments used and the manner in which data was collected. A possible explanation for this could be that these study phases were also the phases in which study participants had the most experience themselves (79% had personal experience in the Study Preparation phase and 69% in the Data Collection phase) and that their personal experiences had made them more positive about experienced levels of achievements in these respected phases due to self-reporting bias (36). This phenomenon could also be an explanation to why the standard with the lowest mean level of achievement was reported for the Dissemination and Communication phase, in which study participants had the least personal experience (38%).

Themes that had been reported to have lower levels of achievement were mainly on the public availability of study protocols, data and publications, proper annotation of the different steps in research and the consideration of stakeholders. These issues of public availability may be considered as issues of open science, for which a push is being made as of recently (37–39). The barriers identified regarding public availability of study

findings and publications seem to have more to do with the form in which they are presented and to whom, rather than to do with unwillingness or being unable to adhere to these criteria. On the other had, public availability of data and the encouragement of reanalyses by local researchers did seem to have a lot to do with the willingness of the GHR community. This seems to be the most striking issue that came forth from the IDIs and was also very much intertwined with other sources of barriers and facilitators. The willingness of stakeholders to provide data was often linked to monetary resources and the incentive to protect one's own interest in this. And the issue of not receiving the proper credit for one's work seems to also be of influence in regards to the public availability of data for reanalyses.

Subgroup analyses seemed to show that women reported lower levels of achievement when it came to the consideration of stakeholders, a theme most related to research fairness. This is in contrast to the overall reported levels of achievement, in which women reported higher levels of achievement than the men did. A possible explanation for this finding could be that women have a more critical view towards fairness than men. A study that compared the perception of experiencing justice in their working environment between men and women found a significant difference in which fairness in procedures seemed to be of greater influence to this perception for women (40). This could attest to the difference found between these groups.
Another difference between subgroups found, were those between researchers with less and more than 3 years of working experience in GHR, in which the former had reported a lower level of achievement in all of the BRIDGE criteria than the latter. Although the number of participants within the latter was quite small (n=5) and no statistically significant results could be found, this finding does correlate to the findings of a study performed, in which the difference in perceptions of research integrity between different ranks among academic researchers in Amsterdam was assessed, that showed that junior researchers had a more negative perception on research integrity climate in comparison to senior researchers (41). On the other hand, another plausible explanation could be that more experienced global health researchers are more likely to comply with the BRIDGE standards and criteria as they have learned, over the years, how to conduct research more responsibly.

Some of the themes identified as possible sources of barriers and facilitators in achieving the BRIDGE criteria had more to do with feasibility than anything else. A lack of resources was often described as a main issue. Whether it concerned time, monetary or staff, all study participants who participated in the IDIs acknowledged to some extent that if they had unlimited resources, they would have been able to reach higher levels of achievement of criteria. This theme also seems to be the centre around which most of the other identified themes revolve. Communication for example, was often referred to as being inadequate due to a lack of funding which leads to a lack of training/briefing. The themes of local context, incentive and ownership are also linked to the resources but unlike communication seem to be more influencing factors on the availability of resources, rather than influenced by a lack of resources.

Two main themes that were identified, which interviewees specifically mentioned as possible sources of facilitators in achieving the BRIDGE criteria, were on communication and knowledge of local context. Interviewees stated that open communication as well as familiarity with local context, were factors for success in their experience in GHR. And

the opposite, being poor communication and a lack of knowledge about local context were specifically mentioned as barriers in achieving the BRIDGE criteria. This was also found in a study on research collaborations between the United Kingdom and Africa, where barriers in establishing and maintaining successful collaborations identified included a lack of knowledge about local context in LMICs and poor communication between research partners from HICs and LMICs (42).

For both the survey and the IDIs, participation was voluntary and participants were assured that their identity would be protected by taking appropriate measures. But a certain degree of social desirability bias (36) cannot be excluded, because of the rather sensitive nature of the subject matter, as researchers are asked to rate research integrity and fairness in their collaboration with research institutions. Therefore it could be possible that study participants reported higher levels of achievement than in reality. And for the IDIs specifically, when interviewees gave more general comments on what they felt should be done in practice rather than what they had experienced, some important issues may have not come to light during this study.

Assessment tool: lessons learned

Survey

The 47 statements of which the questionnaire consisted, were presented per standard in the online survey and every standard was presented on a single webpage, which had to be completed before study participants could move to the next page. This seems to have lead to a considerable amount of study participants dropping out during the questionnaire. When looking at the number of participants who answered the questions per standard (table 3), it becomes apparent that the moments of dropout were in between pages of questions. After the first page of statements (equal to the first standard), 3 participants (8%) dropped out. For each of the next 3 standards, another participant had dropped out. Only between the pages of the fifth and last standard, no participants had dropped out. It was remarkable that no participants dropped out in the middle of standards, indicating that drop out seems to have been linked to the turning of the page. The most plausible explanation that should be considered is that participants were deterred from finishing the survey, as it may have been unclear how much longer the survey would take. Survey fatigue may have played a rather distinct role in the successfulness of data collection in this manner. However, a study performed among undergraduates in the United States, comparing six web based surveys (using SurveyMonkey), showed a similar manner of dropout (being 10% at the start of the survey, decelerating over the length of surveys) (43), in which it was argued that this had less to do with survey fatigue and more with the amount of available information at the start of the survey.

Another possible explanation for this specific pattern of study participant dropout in between pages may be of a more technical sort. In pilot-testing the survey among colleagues, no issues were mentioned. However, no account was taken for different types of web browsers used and it may also be an issue with the quality of internet connections or an issue with the online SurveyMonkey service. This could also explain

the number of participants who did answer the questions on personal characteristics but never started the questionnaire (figure 1) and would also be an alternative explanation for the study findings described above, which was not mentioned in the article (43).

The participant characteristics that were asked for during the survey made subgroup analyses possible. Characteristics such as gender, years of experience in GHR and level of education were easily distinguished by questions with multiple choice answer options. When it came to asking about the location of the home country institute however, this was done by means of an open ended question which made it more difficult to classify these into numeric data during the recoding process. Some participants entered more than one country, which could have made it impossible to reclassify their institute into LMIC or HIC if they would have entered two countries, of which one could be classified as LMIC and the other as HIC. Luckily, in this pilot study, the answers given by those participants who met the inclusion criteria did not result in this difficulty, as the two instances were Netherlands/USA and Singapore/Australia. Seeing as these were all considered to be HIC, this was corrected for by recoding them as such. But in a new study this might become a challenge.

IDIs

During the IDIs it became apparent that there was relatively more emphasis on research fairness than on research integrity, compared to the distribution of these two throughout the criteria in the analytical framework. Almost all of the identified sources of barriers and facilitators in achieving the BRIDGE standards and criteria, found during the IDIs, were related to research fairness. Not only the themes such as ownership and incentive, but also communication mostly had to do with issues concerning collaboration between global health researchers. In the survey findings, the theme of public availability of data and publications were found to have lower mean levels of achievement, which was also something that was widely recognized by the interviewees during the IDIs. It may be possible that these issues were more highlighted by the interviewer because of the lower levels of achievement reported in the survey. However, issues on research integrity were also asked which would often circle back to issues around research fairness as well. This begs the question whether research fairness was appropriately represented in the survey, or if it was perhaps underrepresented in comparison to the number of questions on research integrity.

During the IDIs, it was observed that interviewees would often give more general answers to the questions asked and when specific criteria were discussed it sometimes seemed that it was unclear to interviewees what was exactly meant by some of the statements. Language barriers seemed to have played a role in this, as some English terminology hindered the interviews in some occasions. This is remarkable as interviewees also mentioned language barriers as a barrier to achieving higher levels of achievement of the BRIDGE criteria within the theme of communication. This may have also been an influencing factor during the survey, if study participants misunderstood some of the statements.

Strengths and limitations of this pilot study

In regards to the execution of this pilot study several remarks are to be made. First of all, the response rate in regards to the survey was 115 respondents out of 307 invitees (37.5%). This may be considered average for web surveys as stated by Sammut et al. (44) but it is higher than response rates found in other web-based surveys on research integrity which ranged from 21% to 30% (27,41). The rate of completion in comparison to the number of included study participants, should be considered 30 out of 50 (60%), which falls within the range of the above mentioned surveys (47%-71%) (27,41).

The survey study population showed evenly distributed participant characteristics concerning gender and of institute home countries considered as LMICs and HICs. The distribution of participants with more or less than 3 years of experience in GHR however, showed an underrepresentation of those with more than 3 years of experience. Also, due to the limited number of study participants almost no significant differences were found between the subgroups of study participants. In retrospect, the added value of performing the Wilcoxon rank-sum tests was limited in this study, since the overlapping SD ranges found for the criteria already gave the expectation that no significant findings would be found. However, in a larger study population, this would be expected to have more added value as the power of the study increases and the margin of error decreases (45,46).

Eleven study participants were willing to participate in the IDIs, which was not a large enough group of candidates to successfully perform the purposeful sampling strategy. The planned sample size was to have six interviewees, using the maximum variation strategy. However, only four could be taken. The distribution of the selection criteria concerning gender and mean levels of achievement were evenly distributed throughout the group of interviewees. But this was again not the case for years of experience in GHR, as most of the interviewees had 1-3 years of experience. In order to further explore the difference found based found on years of experience in the survey, ideally the distribution would have even been between less than 3 years of experience and more than 3 years of experience. Also, since all four interviewees had worked for research institutes located in LMICs, a point of view from the HIC in collaboration with the Global South was not represented in the IDIs.

Other limitations of the IDIs lie in the fact that only one IDI was performed per participant. Ideally, the study participants would be interviewed several times to collect more data (47) but due to time restrictions this was not possible. Another limitation to be mentioned is that no saturation could be reached, due to the limited number of participants who could be interviewed. Efforts to strengthen internal validity by means of an additional interview for member checking could have been a partial solution (48) to this. However, due to time restrictions, this was also not performed.

Conclusion and recommendations

The purpose of this pilot study was to make a first attempt into assessing current practices in GHR in relation to the BRIDGE guidelines and to identify possible barriers and facilitators in the compliance with these guidelines. An additional purpose of the study was to develop and evaluate an assessment tool for this assessment. Therefore these will be discussed separately.

Fostering research integrity and research fairness

Trends were identified in regards to the level of achievement of the BRIDGE guidelines. These trends show that in regards to the standards and criteria, some seem to be more achieved in current practices in GHR as experienced by KIT alumni than others. As the issues with the themes regarding research fairness showed both lower levels of achievement in the survey and were more prominently mentioned in the IDIs than for the themes regarding research integrity, it is clear that the development of the BRIDGE guidelines was warranted.

Specifically the criteria on public availability of study protocols, data and publications and proper annotation of the different steps in research seem to be the ones that are most difficult to achieve. Possible sources of barriers in achieving these criteria were found in areas such as understanding of the local context of research settings, incentives and ownership of data of the stakeholders. These issues are to be considered when entering into transnational research collaborations but most of all, these issues show that the scientific global health community should put more efforts towards open science. Initiatives on this have started to present as of recently but I would advocate for more awareness by means of focussing future research endeavours more towards this goal.

Another important theme that was identified, as a possible source of both barriers and facilitators in achieving the BRIDGE standards and criteria, was found in communication. In general, if communication lacked this would hinder research integrity and research fairness. And if communication were actively encouraged, this would facilitate research integrity and research fairness. It would be wise to consider this as an important take home message for all global health researchers.

Recommendations for the RIRF study

As one of the main objectives of this pilot study was to evaluate the means of assessment, on which the RIRF study could build further, recommendations for future research within the field of research integrity and research fairness are to be made.

In conducting this pilot study, issues around research fairness have come more to light than on research integrity in both the survey and the IDIs. Taking this into

consideration, the primary recommendation for the RIRF study is to reassess whether the framework, as it was used in this assessment of the BRIDGE guidelines should be revised more before embarking on a next study, to put more emphasis on research fairness. Therefore, for the planned RIRF survey it might be prudent to have a more evenly distributed amount of questions for issues on both research integrity and research fairness.

Another recommendation for the RIRF survey would be to group some of the main themes that can be derived from the BRIDGE standards and criteria, as the amount of questions in this pilot study may have lead to more dropouts of study participants than necessary. Whether this was due to survey fatigue or technical issues with the online survey, a smaller amount of questions (and therefore with less use of webpages) might lead to a higher rate of completion of the survey in both cases. However, as this could also be due to technical issues and the dropout of study participants was only found in between the pages coinciding with the different standards, it would be advisable to use as little amount of different webpages for the online survey as possible to avoid this dropout as much as possible. Another recommendation concerning the online survey would be to pilot test it in different web browsers as this could also have played a part in the survey dropout. Although this cannot be concluded with any certainty, efforts should be made to reduce the risk of such issues in the RIRF study that is to be performed.

In regards to the questions on participant characteristics, possible issues are to be expected with the use of open-ended questions for determining the context of collaborations between LMICs and HICs. During this pilot study it became apparent that this could lead to study participants declaring to have their home research institute located in more than one country. Therefore, it is recommended that this should be revised into a closed-ended question to avoid these issues. A possible way to anticipate for such difficulties in dividing the study participants into these subgroups would be to ask participants to classify their home research institute themselves into these categories.

References

1. Baker M. 1,500 scientists lift the lid on reproducibility. Nat 2016 5337604. 2016 May 25;452–4.
2. Ioannidis JPA. Why most published research findings are false. Get to Good Res Integr Biomed Sci. 2018;2(8):2–8.
3. Amrhein V, Korner-Nievergelt F, Roth T. The earth is flat (p > 0:05): Significance thresholds and the crisis of unreplicable research. PeerJ. 2017 Jul 7;2017(7):e3544.
4. John LK, Loewenstein G, Prelec D. Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. Psychol Sci. 2012 Apr 16;23(5):524–32.
5. Xie Y, Wang K, Kong Y. Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis. Sci Eng Ethics. 2021 Jun 29;27(4):41.
6. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS One. 2009 May 29;4(5):e5738.
7. Steneck NH. Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions. Sci Eng Ethics. 2006;12(1):53–74.
8. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie Du Sert N, et al. A manifesto for reproducible science. Nat Hum Behav. 2017 Jan 10;1(1):1–9.
9. Academies A-AE. The European code of conduct for research integrity. 2017;
10. Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. The Hong Kong principles for assessing researchers: Fostering research integrity. PLoS Biol. 2020;18(7):1–14.
11. Swaen GMH, Langendam M, Weyler J, Burger H, Siesling S, Atsma WJ, et al. Responsible Epidemiologic Research Practice: a guideline developed by a working group of the Netherlands Epidemiological Society. J Clin Epidemiol. 2018;100:111–9.
12. Koplan JP, Bond TC, Merson MH, Reddy KS, Rodriguez MH, Sewankambo NK, et al. Towards a common definition of global health. Vol. 373, The Lancet. Elsevier; 2009. p. 1993–5.
13. Alba S, Mergenthaler C. Lies, damned lies and epidemiology: Why global health needs good epidemiological practice guidelines. BMJ Glob Heal. 2018;3(5):10–2.
14. Lavery J V, Ijsselmuiden C. The Research Fairness Initiative: Filling a critical gap in global Research Fairness Initiative research ethics. 2018;2:58(0).
15. Musolino N, Lazdins J, Toohey J, IJsselmuiden C. COHRED Fairness Index for international collaborative partnerships. Lancet. 2015;385:1293–4.
16. Walsh A, Brugha R, Byrne E. "The way the country has been carved up by researchers": ethics and power in north-south public health research. Int J Equity Health. 2016 Dec 12;15(1):1–11.
17. Hedt-Gauthier B, Airhihenbuwa CO, Bawah AA, Burke KS, Cherian T, Connelly MT, et al. Academic promotion policies and equity in global health collaborations. Vol. 392, The Lancet. Elsevier; 2018. p. 1607–9.
18. Gautier L, Sieleunou I, Kalolo A. Deconstructing the notion of "global health research partnerships" across Northern and African contexts. BMC Med Ethics. 2018;19(1):49.
19. Rees CA, Ali M, Kisenge R, Ideh RC, Sirna SJ, Britto CD, et al. Where there is no local author: A network bibliometric analysis of authorship parasitism among research conducted in sub-Saharan Africa. BMJ Glob Heal. 2021 Oct 1;6(10):e006982.

20. Cash-Gibson L, Rojas-Gualdrón DF, Pericàs JM, Benach J. Inequalities in global health inequalities research: A 50-year bibliometric analysis (1966-2015). PLoS One. 2018 Jan 1;13(1):e0191901.
21. Lawrence DS, Hirsch LA. Decolonising global health: Transnational research partnerships under the spotlight. Int Health. 2020;12(6):518–23.
22. Costello A, Zumla A. Moving to research partnerships in developing countries. Br Med J. 2000;321(7264):827–9.
23. Addo-Atuah J, Senhaji-Tomza B, Ray D, Basu P, Loh FH (Ellen), Owusu-Daaku F. Global health research partnerships in the context of the Sustainable Development Goals (SDGs). Res Soc Adm Pharm. 2020;16(11):1614–8.
24. Alba S, Verdonck K, Lenglet A, Rumisha SF, Wienia M, Teunissen I, et al. Bridging research integrity and global health epidemiology (BRIDGE) statement: Guidelines for good epidemiological practice. BMJ Glob Heal. 2020;5(10):1–10.
25. Creswell JW. Research design; Qualitative, Quantitative, and Mixed Methods Approaches. 2014.
26. Patton M. Qualitative evaluation and research methods. 1990. 169–186 p.
27. Gopalakrishna G, ter Riet G, Vink G, Stoop I, Wicherts JM, Bouter LM. Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in the Netherlands. PLoS One. 2022;17(2 February):1–16.
28. Huybers T, Greene B, Rohr DH. Academic research integrity: Exploring researchers' perceptions of responsibilities and enablers. Account Res. 2020;27(3):146–77.
29. Cossette P. Research integrity: An exploratory survey of administrative science faculties. J Bus Ethics. 2004;49(3):213–34.
30. Lin GG, Scott JG. Research coordinators experiences with scientific misconduct and research integrity. NIH Public Access. 2012;100(2):130–4.
31. Hofmann B, Bredahl Jensen L, Eriksen MB, Helgesson G, Juth N, Holm S. Research Integrity Among PhD Students at the Faculty of Medicine: A Comparison of Three Scandinavian Universities. J Empir Res Hum Res Ethics. 2020;15(4):320–9.
32. Alba S, Lenglet A, Verdonck K, Roth J, Patil R, Mendoza W, et al. Bridging research integrity and global health epidemiology (BRIDGE) guidelines: Explanation and elaboration. BMJ Glob Heal. 2020;5(10):1–15.
33. Rafidah Zainal, Ayub Md. Som, Nafsiah Mohamed. A Comparative Study on the Implementation Stage of Corporate Integrity in the Malaysian Public Sectors. Reli Rev Ciencias Soc Y Humanidades. 2019;4:123–9.
34. Creswell JW. Educational research; planning, conducting and evaluating quantitative and qualitative research. Pearson; 2012.
35. Mann HB, Whitney DR. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Ann Math Stat. 1947 Aug 9;18(1):50–60.
36. Fadnes LT, Taube A, Tylleskär T. How to identify information bias due to self-reporting in epidemiological research. Internet J Epidemiol. 2009;(January).
37. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. Proc Natl Acad Sci U S A. 2018;115(11):2600–6.
38. Bouter LM, Riet G ter. Replication Research Series-Paper 2 : Empirical research must be replicated before its findings can be trusted. J Clin Epidemiol. 2021;129:188–90.
39. Comisión Europea. H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Eur Comm. 2017;(21 March):1–10.
40. Sweeney PD, McFarlin DB. Process and outcome: Gender differences in the assessment of justice. J Organ Behav. 1997;18(1):83–98.

41. Haven TL, Tijdink JK, Martinson BC, Bouter LM. Perceptions of research integrity climate differ between academic ranks and disciplinary fields: Results from a survey among academic researchers in Amsterdam. PLoS One. 2019;14(1):1–16.
42. Dean L, Njelesani J, Smith H, Bates I. Promoting sustainable research partnerships: A mixed-method evaluation of a United Kingdom-Africa capacity strengthening award scheme. Heal Res Policy Syst. 2015;13(1):1–10.
43. Hoerger M. Participant dropout as a function of survey length in internet-mediated university studies: Implications for study design and voluntary participation in psychological research. Cyberpsychology, Behav Soc Netw. 2010 Dec 1;13(6):697–700.
44. Sammut DR, Griscti DO, Norman PIJ. Strategies to improve response rates to web surveys: A literature review. Vol. 123, International Journal of Nursing Studies. Pergamon; 2021. p. 104058.
45. Fitzner K, Heckinger E. Sample size calculation and power analysis: A quick review. Vol. 36, Diabetes Educator. SAGE PublicationsSage CA: Los Angeles, CA; 2010. p. 701–7.
46. Siegel S. Nonparametric Statistics. Am Stat. 1957;11(3):13–9.
47. Mears CL. In-depth interviews. Research methods and methodologies in education. 2012 Mar 19;19:170-6.
48. Creswell JW, Miller DL. Determining Validity in Qualitative Inquiry. Theory Pract. 2000 Aug 1;39(3):124–30.

Appendix 1 - BRIDGE checklist as adopted from Alba S. et all (24)

#	**Standard 1. Study preparation: carefully prepare the study, in partnership with local researchers, by taking into account existing knowledge and resources and engaging with key stakeholders**
1.1	Plan and execute research in partnership with local researchers. When working in a setting where relevant epidemiological competences are limited or not available, consider what is in the study team's remit to strengthen local capacity
1.2	Identify and engage key stakeholders throughout the study with approaches based on their needs, competences and expectations. Key stakeholders include representatives of affected populations and end-users of research
1.3	Establish the knowledge gap by searching the literature (peer-reviewed publications and grey literature) as well as by consulting (local) experts, representatives of affected populations and end-users
1.4	Develop research questions and objectives in consultation with research partners and expected end-users
1.5	Select study design and research methods to best fulfil the study objectives and give due consideration to multidisciplinary approaches
1.6	Before embarking on primary data collection, assess whether existing data could be used, fully or partly, to fulfill the research objectives
1.7	Ensure data ownership and publication agreements have been agreed by all research partners
1.8	Agree on work plans and governance structures with all study partners. Allocate adequate time, financial and human resources to all phases of the study
#	**Standard 2. Protocol development: prepare a detailed research protocol and ensure it has been approved by relevant ethical review boards if it includes research concerning human participants**
2.1	Prepare a detailed research protocol in consultation with all research partners
2.2	Write a clear and comprehensive analysis section
2.3	Consider studying the effect of locally relevant equity dimensions
2.4	When conducting multidisciplinary research, describe the purpose and strategies to integrate different analytical methods in the protocol
2.5	Strive to make study protocols publicly available, either on a publicly accessible website or in appropriate study registers
2.6	For all data collection and data use concerning human subjects, obtain ethical approval (or a waiver) ideally from all institutions and countries involved in the protocol. In case of multiple review and disagreement, the review of the country where the data are collected should take precedence
2.7	When working in a setting without ethical review boards or review boards with limited epidemiological capacity, consider what is in the study team's remit to strengthen their epidemiological capacity
2.8	Explicitly state any open data access in the protocol submitted for ethical review and in the informed consent documents
#	**Standard 3. Data collection: use valid and reliable instruments and reproducible methods while ensuring culturally appropriate procedures**
3.1	Use valid and reliable research instruments
3.2	Ensure that research instruments are locally adapted and culturally appropriate
3.3	Provide concrete guidance for data collection in a document that is available to all data collection staff
3.4	Select data collection staff according to technical as well as cultural criteria. Clarify the roles and responsibilities for each person involved and provide adequate training and support

3.5	Pilot-test and, if possible, field-test all research instruments prior to the start of effective data collection
3.6	Collect data a respectful and safe manner and in an environment which safeguards the confidentiality of respondents
3.7	Put in place quality assurance and control mechanisms to ensure data accuracy, completeness and coherence
#	**Standard 4. Data management: manage data with reproducible procedures and ensure compliance with relevant data protection rules**
4.1	Put in place data management procedures before effective start of data collection and provide concrete guidance in a document available to all data management staff
4.2	Create and pretest a data entry application prior to effect start of data collection
4.3	Describe all variables in a codebook and consider preparing additional metadata documentation
4.4	Put in place quality assurance and control mechanisms to ensure data accuracy, completeness and coherence
4.5	Annotate all data cleaning and processing steps and strive for reproducibility by means of stored programming code
4.6	For each data file define levels of anonymisation and privacy protection as well as corresponding access rights in line with national and international frameworks
4.7	At the beginning of the study, prepare an electronic secured study file to store all study documentation and outputs. Regularly update this file and archive it the end of the study
4.8	Retain source data safely, in their original form, preserving data confidentiality for as long as has been described in the protocol
#	**Standard 5. Data analyses: analyse data according to the protocol and integrate statistical analyses with approaches from other disciplines in the study**
5.1	Only work with personal identifiers that are necessary to answer the research questions
5.2	Conduct statistical analyses in accordance with the protocol and distinguish preplanned from exploratory analyses
5.3	Fully annotate all analysis steps and strive for reproducibility by means of programming code
5.4	In multidisciplinary studies, integrate statistical analyses with analyses from other study disciplines in an iterative process to coherently address the research objectives
5.5	Put in place quality assurance and quality control mechanisms to ensure that data has been correctly analysed
#	**Standard 6. Dissemination and communication: report and disseminate results, preferably in the public domain, with means of communication which appropriately target key stakeholders**
6.1	Develop user-specific dissemination and communication plans in consultation with key stakeholders (representatives of the affected populations and end-users)
6.2	Report data in a non-stigmatising, non-discriminatory, culturally sensitive and non-identifying manner
6.3	Conform to reporting guidelines for the given study design and methods in academic publications
6.4	Put in place quality assurance and quality control mechanisms to ensure complete, accurate, accessible and interpretable data reporting
6.5	Consider indexed open access journals for scientific publications
6.6	On study completion, consider publication of the archive in an openly accessible online repository. Consult key stakeholders and research partners to identify strategies within the study team's remit to encourage as much as possible reanalyses by local researchers

Appendix 2 – *Percentage of respondents who answered "Don't know/not applicable"*

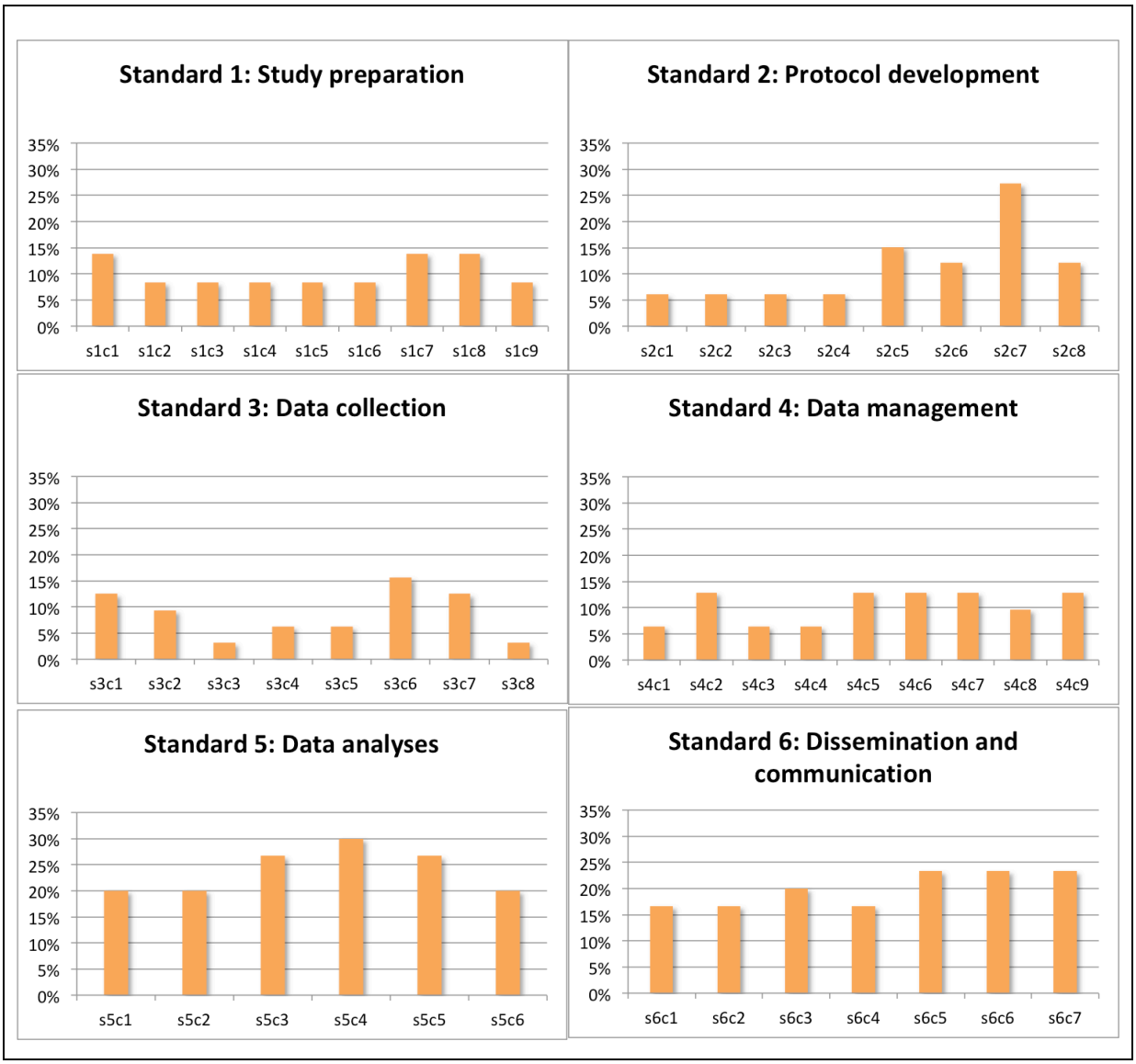

Figure. Percentage of respondents who answered "Don't know/not applicable".
Percentage of respondents who answered "Don't know/not applicable" per standard, for respectively the Study preparation phase (n=36), the Protocol development phase (n=33), the Data collection phase (n=32), the Data management phase (n=31), the Data analyses phase (n=30) and the Dissemination and communication phase (n=30).