RESEARCH ARTICLE

Evaluation of a guideline developed to reduce HIV-related stigma and discrimination in healthcare settings and establishing consensus

Garumma Tolu Feyissa[1,2,3]*, Craig Lockwood[3], Mirkuzie Woldie[1,2,4], Zachary Munn[3]

1 Jimma University, Department of Health, Behavior and Society, Jimma, Ethiopia, **2** Ethiopian Evidence Based Health Care Centre: JBI Center of Excellence, Jimma University, Jimma, Ethiopia, **3** The Joanna Briggs Institute, the University of Adelaide, Adelaide, Australia, **4** Department of Health Economics, Management and Policy, Jimma University, Jimma, Ethiopia

* garummatolu@yahoo.com

Abstract

Background

Developing guidelines and policies is critical to address HIV-related stigma and discrimination (SAD) in healthcare settings. To this end, a multidisciplinary panel developed a guideline to reduce SAD. This project evaluated the appropriateness of implementing the guideline in the Ethiopian context.

Methods

A consensus of the expert panel was established through a modified Delphi technique which was followed by a panel meeting. Initial tentative recommendations were distributed to experts through e-mails to be evaluated using the modified guideline implementability appraisal (GLIA) v.2.0 checklist.

Results

In the first round of the Delphi survey, all (13) panel members evaluated the guideline. The overall score for the general domain of the modified GLIA checklist was 96.56%. The scores for individual recommendations ranged from 68.33% to 92.76%. Maximum and minimum scores were attained for measurability (97.71%) and flexibility (59.77%) domains respectively. Percentages mean score lower than 75% was obtained for flexibility and validity domains. Participants suggested that additional tools and training should be added to the guideline. In the second round of the survey, all the recommendations received endorsement with scores above 75%. Maximum and minimum scores were attained for measurability (100%) and flexibility (86.88%) domains respectively. During the panel meeting, issues of responsibility for implementing the guideline were discussed.

Conclusion

The project evaluated implementability of a guideline developed to reduce HIV-related SAD in healthcare settings. The Delphi survey was followed by a half-day meeting that helped in further clarification of points.

Background

People living with the human immunodeficiency virus (HIV) are confronted with the physical, psychological and social impacts of the disease [1–5]. Stigma and discrimination (SAD), also called the "third phase of HIV/AIDS epidemics", have been among the obstacles challenging actors working on the prevention and control of HIV [6]. In healthcare settings, SAD related to HIV are manifested in various forms such as: differential care or refusal to treat, testing and disclosure of the sero-status of clients without consent, verbal abuses or gossip, marking the files of patients, isolating them and excess use of precautions [7, 8].

The limited awareness of SAD, how they manifest and their consequences, prejudicial and stereotypical attitudes related to gender identity and sexual activity, and fear of HIV transmission are among factors contributing to SAD in healthcare facilities [9]. Hence, developing appropriate guidelines, policies, and redress systems and appropriate orientation of the rights and responsibilities of HCWs and patients are critical [9]. Cognizant of this, we have systematically developed a list of working recommendations to reduce SAD in healthcare facilities. Systematically developed guidelines are the source of summarized information [10]. Nevertheless, the development of guideline recommendations by itself is not enough. Other factors such as environmental and contextual factors need to be considered before making final decisions on the implementation of the guideline [11, 12].

Factors such as reviewing, reporting and publishing guidelines have been found to enhance the implementation of the guidelines [13]. On the other hand, Jordan et al. argue that dissemination should involve an active process apart from the mere publication of guidelines [10]. Moreover, before officially publishing or disseminating a guideline, internal and external evaluation is required to promote the uptake of the guideline [14, 15]. In addition to the development of tools to assess the rigor of the guideline development process [16], researchers have developed tools that help to assess both the rigor and implementability of guidelines [17]. Guideline developers and experts recommend assessing recommendations included in practice guidelines using guideline implementability checklists to make sure that the recommendations are clear and easy to implement [14, 15].

We have developed guideline recommendations based on an analysis of global evidence retrieved through literature searching. Therefore, this project aimed to assess the clarity, acceptability, implementability and relevance of the current guideline using Guideline Implementability Appraisal (GLIA version 2.0) checklist [18].

The objective of this project was to evaluate the appropriateness of the guideline developed to reduce HIV-related S&D to be implemented in the Ethiopian context. Specifically, the project aimed:

- To evaluate the appropriateness of the guideline to the Ethiopian context through a multi-round of Delphi surveys among the guideline panel.

- To evaluate the appropriateness of the guideline through a survey of external experts.

- To make amendments to each recommendation included in the guideline based on the comments of the experts.

Methods

This project assessed the drafted recommendations for feasibility and appropriateness to the Ethiopian context. Consensus of the experts engaged in the evaluation was established through a modified Delphi technique [19].

Rationale for the use of the Delphi technique in this project

The Delphi technique involves a series of questionnaires that are used to test opinion consensus amongst a group of experts [20, 21]. The technique can be conducted by email, online surveys or by post [21]. It is a preferable method of choice when there is little evidence regarding the topic, when participant anonymity is required, and when the cost and practicalities of bringing the participants together is prohibitive [22]. By assuring anonymity, it reduces the effect of dominant individuals and unwillingness to abandon publicly expressed opinions [23]. The Delphi technique also reduces reluctance to mention opinions that are unpopular, disagree with one's associates, modify previously stated positions [24].

The choice for the specific type of consensus method is determined by the purpose of the study, the availability of scientific evidence in the field, the model of participant interaction, time and costs [22]. The aim of the current project was to translate research evidence into practice through the development of an evidence-informed guideline based on the consensus of experts. The development of a guideline needs a rigorous process to achieve consensus of experts. In such circumstances, the Delphi technique is supposed to be more suitable compared to other consensus building methods. In addition, the Delphi technique gives adequate time to the experts to exhaust options before making decisions [25]. Moreover, the current project was aimed to seek the opinion of experts by keeping their responses anonymous and allowing them to freely express their opinions through e-mail surveys. Hence, we selected the modified Delphi technique compared to other consensus development techniques. In the current project, the Delphi technique was modified by complementing it by other technique (face-to-face meeting) and by collecting date through e-mail.

The Delphi technique is a hybrid of qualitative and quantitative methods [26]. The potential limitation of the Delphi technique is that some people who participate in early rounds may drop out in subsequent rounds. However, in the Delphi technique, it is possible to involve large numbers of participants from different geographical areas [20]. In addition, anonymity can help to increase the response rates in the Delphi surveys [23]. Moreover, experts participate in the consensus process at a stage when convenient to them and only contribute to those aspects that they feel best able to contribute [27]. The Delphi technique also reduces reluctance to mention opinions that are unpopular, disagree with one's associates, modify previously stated positions [24].

The other potential limitation of a modified Delphi approach is the absence of face-to-face engagement with panel members [28]. In the current project, this limitation was minimized by incorporating two panel meeting sessions, one before the start of the modified Delphi survey and one after the second round Delphi survey. This has helped to clarify and discuss vague points.

The Delphi technique has been used in health disciplines since the 1970s [20]. It has been used by researchers to translate scientific knowledge and professional experience into informed judgment, in order to support effective decision-making [29]. The Delphi technique

has been reported to be the most widely used consensus method for developing clinical guidelines [30–32]. Delphi techniques have been used to develop guidelines, to establish consensus on the use of the guidelines and to establish and evaluate how well a clinical practice is conforming to guidelines [33]. In the current project, the Delphi technique was used to establish consensus on the use of the each of the recommendations that constituted a guideline to reduce HIV-related stigma and discrimination in Ethiopian healthcare settings.

Delphi process

The Delphi procedure starts with the selection of experts and is executed in a series of rounds [22, 23]. In a Delphi survey, appropriate selection of experts is essential for ensuring the quality of the data and increasing response rates. There is no standard definition of expert [25] and the definition depends on the specific objective of the research [25], but in general an expert is someone who has some knowledge of a specific subject [25, 34]. In the current project, experts were people who were knowledgeable of the subject matter by virtue of their role as clinicians with HIV patients, managers for HIV programs or researching on HIV. Experts may be selected based on records of relevant publications, their relationship with the topic and institutional positions they hold [27]. In the current project, the judgment of expertise was made based on their contribution in the field. Hence, researchers with relevant research projects and publications; health service managers and health professionals working on clinical or programmatic areas of HIV were selected as members of the guideline working group and experts for the current Delphi study. All the experts selected were based in Ethiopia and had adequate knowledge of the study context. Apart from their expertise, the availability and commitments of the experts in the field were considered in selecting the panel members. The snow balling method was used to identify the experts. Finally, experts who were willing to participate were included in the multi-round survey.

Panel size

There is no consensus on the panel size required for Delphi studies [29, 35]. Different Delphi studies have used different sample sizes ranging from as small as five to as large as 2865 [29]. In the Delphi technique, sample size does not depend on statistical calculations; rather it depends on the dynamics of arriving at consensus [36]. Some experts in Delphi techniques recommend careful selection of the panel for the specific topic of interest instead of increasing sample size or making the sampling process random [37]. In this project, 13 experts accepted our invitation and participated in the survey.

As a facilitator of the Delphi technique, the principal investigator set deadlines for each round of the Delphi and he used e-mail reminders for non-responders as an additional mechanism for increasing the response rate. The principal investigator sent the e-mail reminders three days after the deadline [27]. Respondents were given a three-week period for each round of Delphi [27]. As in other Delphi techniques, the opinion of every group member was reflected in the final group response [24]. The statistical average of the final opinions of the individual members was used to define group opinion [24].

Data collection

After obtaining the list of experts, the principal investigator (GTF) made initial contacts to all experts giving them the purpose and procedures involved in the project and requesting them to participate in the development of the guideline. After receiving consent, the principal investigator sent the experts initial tentative recommendations by e-mail. Experts were asked to

comment on each recommendation. We analyzed and summarized both qualitative and quantitative responses [22, 23].

There are three options to start Delphi round one. The first option is where Delphi round one is conducted as a qualitative study using open-ended questions to develop quantitative tools for the successive rounds [38]. In this approach, the first round is used to identify issues to be addressed in later rounds. The second option is where qualitative data can be collected through focus groups or interviews before the Delphi study and used to inform a quantitative first round of the Delphi [25]. The third option is where the quantitative first round is informed through a literature review or clinical practice [25, 39]. The first approach is often used in a classical (original) Delphi [40]. The second and third approaches are usually used in a modified Delphi technique [40].

In the current project, the tentative recommendations were informed by a systematic literature search and content analysis of the evidence. In this project, the modified Delphi, sometimes called 'e-delphi'[40] was used. The purpose of the modified Delphi technique in this project was to get a consensus among the guideline panel on the tentative recommendations, and to modify the recommendations based on the responses of the experts. Therefore, the third approach was employed. Hence, experts were asked to rate each tentative recommendation using the Guideline Implementability Appraisal (GLIA V.2.0) checklist [18]. The GLIA checklist has options for both close-ended responses and open-ended responses. Hence, in addition to rating the recommendations, the panelists were asked to provide their suggestions on how to improve the implementations, feasibility and/or wordings of the specific recommendations. Participants were also encouraged to comment on the main guideline using track changes and highlights. The GLIA v.2.0 checklist was modified and used to assess the implementability of the guideline [18]. The GLIA v.2.0 [18] instrument contains 30 items in nine domains: global quality, executability, decidability, validity, flexibility, effect on process of care, measurability, novelty and computability [18]. Out of these, the last domain (computability) is used when there is a plan for electronic implementation [18]. Since this will not be part of the current work, the four items in this domain were not included in the questionnaire.

In this project, a modified GLIA v.2.0 checklist was used to assess the implementability of the guideline. The comments provided by the experts were incorporated into the successive round of the Delphi. In the subsequent round of the Delphi, we asked participants whether they would agree with the modified recommendations [22, 23]. We sent additional ideas in each round of the Delphi to the experts in the respective subsequent rounds [22, 23].

There is no template indicating the exact number of rounds needed for a Delphi study. Such decisions are pragmatically made by the researcher. Hence, the procedure is reiterated until the stability of responses is achieved [23]. Stability of responses is defined as "the consistency of responses between successive rounds of a study"[41](pp.84). Dajani et al. recommends measuring the level of agreement only if a stable answer is reached [41]. For each recommendation, once stability of the responses is achieved, consensus will be established [41]. For this project, we did not employ statistical calculation to measure stability. This is because, some recommendations were merged, and additional new recommendations were added and dropped iteratively, it was not practical to employ statistical techniques such as weighted kappa. In addition, after two-rounds of the Delphi surveys, we found no new comments from the panel and the panel decided that there should not be additional survey. Therefore, a half-day face-to-face meeting was conducted where the panel members were asked whether their comments were addressed, and other practical considerations related to the guideline.

For this project, recommendations having a general agreement of 75% and above were incorporated into the guideline. Recommendations with a rating lower than 75% were considered for modification to be incorporated into the subsequent rounds based on the comments

of the respondents [23, 42]. In addition, specific comments given for each recommendation were considered for making modifications, adding or dropping a recommendation. The Delphi series stopped after ensuring that there was no newer comments emerging and a minimum of 75% level of agreement was attained for each recommendation. The Delphi process is normally expected to achieve both consensus and non-consensus [42]. Therefore, in the current project, recommendations for which experts consistently disagreed were excluded or modified.

Data quality control

In Delphi techniques, the opinion of every group member is reflected in the final group response [24]. Since decisions are made based on opinions of groups in the real world, Delphi techniques are believed to provide evidence of face validity [43]. In addition, Delphi is conducted in successive rounds, contributing to concurrent validity of the findings [44]. Researchers also believe that a Delphi technique provides reliable findings, because it achieves interaction among experts and at the same time avoids individual influences. Delphi overlaps both interpretive/qualitative and positivist/quantitative paradigms. Hence, researchers recommend the use of the term 'trustworthiness' to establish rigor in a Delphi study.[45] The concept 'trustworthiness' encompasses credibility, transferability, dependability and confirmability [46].

In Delphi studies, credibility is established by ongoing iteration and feedback given to the experts [47]. Therefore, the very beginning of the Delphi process makes it credible. In this project, dependability was enhanced by including relevant experts in the field.[48]. Confirmability is achieved through the collection of thick descriptive data, negative case analysis and arranging for a confirmability audit and establishing referential adequacy [46]. In this project, we kept accurate records of participants' comments and responses in each round. We sent the comments of experts to the panelists in subsequent rounds. In addition, there was a face-to-face meeting prepared for further clarification. The transferability of an evidence is based on the similarity of contextual factors in the settings [49]. Therefore, other researchers and guideline implementers or developers were advised to take the consideration of the similarities of their respective contexts with the current situation and the current context of Jimma University Medical Centre (JUMC) when considering the potential transfer of the evidence into other settings.

Data analyses

We conducted qualitative content analysis [50, 51], the comments and we used the result of the analysis to modify the recommendations. In addition, we conducted the following quantitative analyses:

1. Percentage response rates,

2. Percentage scores for each domain of GLIA V.2.0: the total score for each GLIA domain was calculated by summing up total scores for all panel members. Then, the percentage score was obtained by dividing the total score by the maximum possible score.

3. Percentage agreement for each recommendation was calculated for each round of the Delphi. This information was used to modify recommendations, especially those with endorsement of less than 75%. In the cases where experts did not describe reasons for non-endorsement and for controversial issues, discussions on the recommendations were made through face-to-face meetings amongst the panel.

In this project, we wanted to take into consideration the input of each member of the panel. Instead of taking individual responses as outliers and rejecting them, a mechanism was in place in which they would clarify their opinions, which opens up for further comment by other members of the panel. Moreover, the panel consensus data were complemented with external panel review.

Ethical approval and consent to participate

The project has ethical approval both from the Institutional Review Board (IRB) of Jimma Institute of Health (JIH) at Jimma University (RPGC/389/2016) and the University of Adelaide Office of Research Ethics, Compliance and Integrity (ORECI) (approval number H-2016-140). Prior to the data collection, the objective of the research, potential harms and benefits of participating in the project were described to participants. Participants were provided with complaints procedure and information sheets, based on which informed consent was obtained. Anonymity of responses was assured by not disclosing the identity of participants.

Results

A formal consensus was sought from all the panel members using two rounds of panel surveys and an external panel review. This section describes results of these surveys.

First round Delphi survey

In the first round of the Delphi survey, all (13) panel members evaluated the guideline. The overall score for the general domain of the GLIA version 2.0 score was 112 (% of maximum possible score = 95.73%). Maximum score was achieved for the measurability domain (96.65%) and the minimum score was recorded for the flexibility domain (59.97%). A percentage mean score lower than 75% was obtained only for two domains: flexibility and validity domains (Table 1). The experts provided comments on how to improve or why modifications were needed for individual recommendations included in the guideline. The comments given were categorized into:

1. **General comments:** Comments that were provided for the entire guideline. These comments were suggestions for additional tools and training that should be part of the guideline

Table 1. Guideline implementability (GLIA V.2.0) domain scores.

GLIA Domain	Internal evaluation						External evaluation		
	Round 1			Round 2					
	Mean	SD	%age score	Mean	SD	%age score	Mean	SD	%age score
Executability	21.81	3.35	83.88	15.38	0.77	96.13	10.50	1.51	87.50
Decidability	33.1	3.36	84.89	23.85	0.38	99.38	17.25	1.54	95.83
Validity	17.48	4.77	67.23	15.85	0.554	99.06	10.50	1.93	87.5
Flexibility	23.39	4.23	59.97	20.85	0.38	86.88	9.66	1.37	53.70
Effect on process of care	24.71	1.04	95.04	15.77	0.44	98.56	11.33	0.89	94.44
Measurability	25.13	0.96	96.65	16.00	0.00	100	10.17	0.39	84.72
Novelty	33.44	2.14	85.74	23.08	0.49	96.08	17.00	1.41	94.44

NB: GLIA: Guideline Implementability Appraisal, SD: standard deviation, %age score: percentage score

2. **Comments on specific recommendations:** Comments questioning the clarity and feasibility of implementing the recommendations.

The scores for individual recommendations ranged from 151 (68.33%) to 205(92.76). Six recommendations received an endorsement of lower than 75%. The recommendations with endorsement lower than 75% were:

1. **Counselling and behaviour change programs to address self-stigma (endorsement score = 71.04%)**
   The most important reasons for the low score for this recommendation was described as lack of detailed description of the recommendations and failure to specify the type of behavioural change programs.

2. **Group intervention through telephone support for people living with HIV (endorsement score = 68.33%)**
   The feasibility of this intervention was questioned by the panel. Therefore, this recommendation was brought for panel discussion during the second round panel meeting.

3. **Micro-finance and livelihood programs to create economic opportunities (endorsement score = 70.14%)**
   Concern was raised because participants claimed that it was not the mandate of healthcare institutions to provide microfinance interventions and resource-wise, this recommendation was reported to be not feasible. Therefore, this recommendation was brought for panel discussion during the second round panel meeting.

4. **Training programs to gain facilitation skills, processes to collect and analyse data for advocacy (endorsement score = 70.14%)**
   This recommendation was rated a low score because of limited description linked with it. The feasibility of the recommendation was also questioned.

5. **Developing stigma and discrimination reduction policies with employees (endorsement score = 73.76%)**
   The panel requested description of this recommendation, specifically by linking with previous research findings.

6. **Programs, offices and institutions need to advocate temporary special measures such as affirmative action for women and special forums for participation (endorsement score = 71.04%)**
   The feasibility of this recommendation was questioned as it was perceived by some panel members to be beyond the scope of health institutions. In addition to the above comments targeting individual recommendations, as mentioned in Table 2, the panel suggested that some recommendations should be merged. The main comments made by the panel during the first round survey are summarized in Table 2. Based on the first round comments, modifications were made. The second round survey was then conducted after incorporating comments from the first round and modifications to the guideline.

Second round Delphi survey

Eight of the 13 (61.5%) panel members responded to the second round survey using the GLIA V.2.0 checklist. Five panel members did not provide ratings during the second round Delphi survey. Of these, four of them participated in the second round panel meeting. In the second

Table 2. Summary of comments provided during first round survey.

S/n	Comments	Actions/resolution
General comments		
1.	The sequence of applying these recommendations is not clearly documented	This has been indicated at the end of the recommendations incorporating steps in implementation
2.	In the introduction part, the intended audience should also include non-health disciplines such as psychology and sociology who work to improve the psychosocial well-being of PLHIV	Accepted
3.	For most of the recommendations: patient characteristics (co-morbidities) were not mentioned	Most recommendations work for all types of HIV patients regardless of their co-morbidities
4.	Settings such as faith-based organizations may be included as part of the guideline	This is beyond the scope of the current guideline, which is limited to healthcare settings
5.	The guideline should be broad, and the scope should be beyond the health sector	This cannot be addressed within the time frame. After this project is over, we may consider developing guidelines for other settings
6.	Additional tools should be part of the guideline	Accepted and added tools to be posted and tools for monitoring and evaluation
7.	Key population should be defined	Accepted
8.	People associated with the virus should be defined	Accepted
9.	Stigma occurs when those health care workers who are not aware of HIV-related stigma provide services to HIV patients. Therefore, the type and role of service providers needs to be specified.	Brought for discussion by the panel during the second meeting and further explored during key informant interviews
Comments on specific recommendations		
10.	RN1.4, RN2.4 and RN3.4 are fragmented and can be better strengthened if they are merged together.	Accepted and merged the recommendations So, RN1.4, RN2.4 and RN3.4 were merged
11.	RN2.1 should be supported with evidence	Accepted, reference and quality of evidence included
12.	RN2.1 and RN2.2 can be merged	Accepted
13.	RN 2.2. is not detailed	Accepted
14.	RN2.3 is not detailed. Group support through telephone is not clear enough. Are you going to call them or text them through SMS? It is not feasible, and the quality of evidence is also very low. Also, it is better to use references	Brought for discussion by the panel during the second meeting
15.	One of the recommendations, micro-finance interventions is not feasible	Brought for discussion by the panel during the second meeting
16.	RN2.4 needs resources	Suggestions will be sought from panel members on whether the allocation of such resources is feasible will be discussed
17.	RN2.6 is not specific	The recommendation was dropped
18.	RN2.6 is difficult to measure unless we put measurement parameters	Accepted
19.	RN3.1 is not detailed	Accepted
20.	RN4.2 is not detailed	Accepted
21.	RN6.2 Needs details	Accepted
22.	RN6.3 needs details	Accepted
23.	RN1.4 is not feasible	Accepted
24.	RN4.1. is not feasible	Accepted
25.	Co-morbid mental illness among HIV clients plays critical role in worsening the stigma towards HIV patient. So, consider mental illness	This will broaden our scope. We may consider another guideline for this.

(*Continued*)

Table 2. (Continued)

S/n	Comments	Actions/resolution
26.	All recommendation need at least orientation and training	The guideline will be introduced through different methods including orientation and training. And additional methods will be further sought from the panel
27.	RN1.2, RN 6.3 and 6.11 need to be merged or be described using a single recommendation.	Accepted

NB: HIV: Human immunodeficiency virus, PLHIV: People Living with HIV, RN: Recommendation number

https://doi.org/10.1371/journal.pone.0198781.t002

round panel meeting, all the comments in the first round and second round were summarized and discussed. Hence, those members who missed the second round survey got the opportunity to reflect on their ideas in the meeting. In the second round, the general domain received an endorsement score of 64/72 (88.89%). Maximum score was attained for the measurability domain (100%). A minimum score was recorded for the flexibility domain (86.88%) (Table 1). In the second round of the Delphi survey, each recommendation received an endorsement of over 75%. Only a few comments were raised by the panel. The summary of the comments and the respective resolutions made following the comments is shown in Table 3.

The highest percentage mean score was attained for the measurability domain (100%) and the lowest mean score percentage was attained for the flexibility domain (88.88%) (Table 1). All individual recommendations received endorsements with scores over 75%. Since there were few comments given in the second round survey and the ratings for the recommendations were also high, the panel decided not to have additional surveys. Instead, a second round panel meeting was called to discuss in person the comments made thus far and the modifications made. Further comments were sought from the panel. Major points raised during the meeting are briefly presented below.

Major points of discussion during the second round guideline panel meeting

1. **The responsible body for implementation of the interventions should be clearly specified**:
   Based on detailed discussions, the panel resolved that all health professionals, healthcare facility administration and HIV prevention and control offices are responsible for the interventions in the recommendations be included in the guidelines. The panel recommended that training should be provided for those PLHIV who provide psychosocial support, adherence support and peer support for PLHIV.

Table 3. Summary of comments made during second round and the respective resolutions.

S/n	Comments	Actions/resolution
1.	Details of peer education intervention is not presented	Accepted
2.	Who is responsible for implementing the recommendations? HAPCO or Hospital?	To be discussed during panel meeting
3.	For RN2.0, include the term expert patients to describe patients involved as service providers. This will match with the context	Accepted

NB: RN: Recommendation number, HAPCO: HIV Prevention and Control Office

https://doi.org/10.1371/journal.pone.0198781.t003

2. **Whether microfinance intervention can still be part of the guideline:**
   The panel decided that HIV Prevention and Control Office (HAPCO) and healthcare facilities can routinely link patients to support organizations. Nevertheless, they agreed that it is very difficult for them to provide financial interventions, such as microfinance interventions.

3. **Whether telephone support interventions are still feasible for the context:**
   The panel resolved with the consensus that in the Ethiopian context, there is no adequate evidence indicating that such interventions are feasible. However, they all agreed that these interventions (phone calls and reminder texts) can be included as alternative methods for the provision of psychosocial support.

4. **Who is responsible for informing the rights and responsibilities to patients?**
   The panel resolved with the consensus that all health professionals should routinely inform patients about the details of procedures, their rights and responsibilities. In addition, healthcare facility administration and HIV Prevention and Control Office (HAPCO) are responsible to make sure that information is provided to patients on their rights and responsibilities. This information should include the rights that each patient has regardless of his or her sex, disease status, age and other characteristics.

5. **Whether translating the guideline into local language is needed:**
   The panel decided that for healthcare professionals, there is no need to translate the guideline into local languages. Nevertheless, the training manual that may be prepared in the future for peer supporters and expert patients (non-professionals) should be translated into local languages.

6. **Arrangement of recommendations**
   The panel suggested that the recommendations should be arranged, not under guiding principles, but under major thematic areas.

Evaluation by external experts

Of the 13 experts invited to participate in the evaluation, six agreed to evaluate the guideline using the same checklist that internal evaluators used. The external experts gave an overall score of 51 (94.44%) to the general domain of GLIA. Each recommendation received an endorsement over 75%. The maximum score was recorded for the decidability domain (95.83%) and minimum score was attained for the flexibility domain (53.70%). The external panels did not provide many comments. Major comments made were categorized under general comments, comments specific to individual recommendations and comments related to format of the guideline (Table 4).

Discussion

This project attempted to evaluate a guideline developed to reduce HIV-related stigma and discrimination using guideline implementability appraisal (GLIA) version 2.0 checklist. The internal evaluation was conducted using two rounds of the Delphi survey that was followed by a face-to-face meeting of the guideline panel. The Delphi surveys were complemented by an additional evaluation by external experts.

In the first round Delphi survey, a percentage mean score lower than 75% was obtained for two domains: flexibility and validity domains of GLIA V2.0 checklist. This indicated that more work was needed with including detailed descriptions on areas such as strength and quality of

Table 4. Summarised comments from the external panel.

S/N	Comments	Resolution/actions
General comments		
1.	Recommendations should be action-oriented rather than descriptive. Some recommendations are not identifiable because of long descriptions	Accepted
2.	Settings in which the guideline is to be implemented is not clearly described	Accepted
3.	The guideline mainly focuses on the provider or user of the guideline and simply highlights the target. The targets must be described in detail in a separate section.	Accepted
4.	Target organizations for the guideline are not mentioned except on the cover page.	Accepted
5.	The required service modifications are not mentioned	Accepted
Comments related to the format of the guideline		
6.	Boxes for strategies and recommendations need to be separate	Accepted
7.	Indicators need to be presented clearly for recommendations	Accepted
8.	It is better to put boxes and tables at the end of description rather than putting them in the middle of text descriptions	Accepted
Comments on specific recommendations		
9.	RN3.3 does not show how opinion leaders execute their jobs	Accepted
10.	RN43 does not detail how to empower PLHIV	Accepted
11.	For RN61, RN62, RN63 AND RN64, strategies for implementation was not addressed well	Accepted
12.	RN33 does not show logical sequences	Accepted
13.	RN 51 is not detailed	Accepted
14.	No evidence presented for RN61, RN62	Accepted

NB: PLHIV: People Living with HIV, RN: Recommendation number

https://doi.org/10.1371/journal.pone.0198781.t004

recommendations and detailed justifications of recommendations. Therefore, modifications were made before sending the guideline for the second round evaluation. The modifications made were: incorporating strength and quality of recommendations for those recommendations for which such data were available.

As the experts involved in the Delphi survey were also members of the guideline working group, it was my expectation that the risk of dropping out from the study would be minimal. Nevertheless, in the second round, we obtained a response rate of 61.5%, which was lower than our expectation. This is, however, an expected limitation of Delphi techniques [20]. In addition, it is a common obstacle that guideline developers face when using the GLIA check-lists as it is a long instrument and may result in low response rates [52]. However, the instrument provides an opportunity for a comprehensive evaluation of guideline recommendations. It helps to assess both implementability and rigor of recommendations [52].

The other potential reason for delayed responses and low response rates in the current project might be because the experts were occupied with other tasks and that the current project was conducted within a tight schedule. We had made efforts to reduce delays and drop outs by setting deadlines, e-mail and telephone reminders. Such mechanisms have also been used by previous researchers employing Delphi techniques [27]. On the other hand, the same experts who failed to provide responses for the second round survey participated in a panel meeting where they got an opportunity to reflect on their opinions. In the panel meeting, a summary of the comments and modifications made in all rounds were presented and reflections were made by all participants. Hence, the attrition bias related to drop outs was minimal.

During the external panel survey, the lowest score was recorded for the flexibility domain (53.70%). This was an indication that notified us to make the emphasis on the quality and

strength of recommendations. This was a partially expected response as some recommendations still lacked quality and strength of evidence supporting them. Hence, for such recommendations, we indicated them as 'no quality of evidence assigned'. Later, some of such recommendations were assigned as good practice points.

In addition, there was a concern by external reviewers regarding feasibility issues. Some enquired about the commitment of Jimma University Medical Center (JUMC) for availing continuous supply of materials for standard precautions. Therefore, this was later explored in detail during the key informant interviews (unpublished report). On the other hand, the response 'not applicable (NA)' for question 18 might have contributed to the low score in the flexibility domain. The question enquires whether the recommendations were made with the consideration of co-morbidities among clients, which was not practical for the current guideline.

In general, except assigning a low endorsement score for the flexibility domain, the external panel endorsed all individual recommendations with scores above 75%. For the current Delphi survey, since some recommendations were merged, and additional new recommendations were added and dropped iteratively, it was not practical to employ statistical techniques, such as weighted kappa, index of predicted association and McNemar chi-square tests [41, 53–55] to measure stability and correlation of response rates between rounds. Nevertheless, additional comments were not forthcoming during the second round and during external expert evaluations. In addition, during the second panel meeting, a detailed discussion was held both on the comments and the modifications made to address the comments.

The current project employed a modified Delphi technique to establish consensus on recommendations based on the best available evidence from systematic reviews. Such techniques have been used by previous researchers to develop guidelines [28, 56]. One of the potential limitations of a modified Delphi approach is the absence of face-to-face engagement with panel members [28]. In the current project, this limitation was minimized by incorporating two panel meeting sessions, one before the start of the Delphi survey and one after the second round Delphi survey. This has helped to clarify and discuss vague points. However, before implementing the guideline, it is critical to identify contextual and environmental factors to tailor the implementation of the guideline to local context.

Conclusion

The current project evaluated the implementability of a guideline developed to reduce HIV-related stigma and discrimination in healthcare settings. The project employed both internal and external evaluation. The Delphi survey was followed by a face-to-face meeting that helped in further clarifications of points and addressing some of the limitations of the series of the Delphi surveys.

Acknowledgments

Author Contributions

Conceptualization: Garumma Tolu Feyissa, Craig Lockwood, Mirkuzie Woldie, Zachary Munn.

Data curation: Garumma Tolu Feyissa, Craig Lockwood, Mirkuzie Woldie, Zachary Munn.

Formal analysis: Garumma Tolu Feyissa.

Funding acquisition: Garumma Tolu Feyissa.

Investigation: Garumma Tolu Feyissa.

Methodology: Garumma Tolu Feyissa, Craig Lockwood, Mirkuzie Woldie, Zachary Munn.

Project administration: Garumma Tolu Feyissa.

Resources: Garumma Tolu Feyissa.

Software: Garumma Tolu Feyissa.

Supervision: Garumma Tolu Feyissa.

Validation: Garumma Tolu Feyissa, Craig Lockwood, Mirkuzie Woldie, Zachary Munn.

Visualization: Garumma Tolu Feyissa, Craig Lockwood, Mirkuzie Woldie, Zachary Munn.

Writing – original draft: Garumma Tolu Feyissa.

Writing – review & editing: Garumma Tolu Feyissa, Craig Lockwood, Mirkuzie Woldie, Zachary Munn.

References

1. Catalan J, Harding R, Sibley E, Clucas C, Croome N, Sherr L. HIV infection and mental health: Suicidal behaviour—Systematic review. Psychology, Health & Medicine. 2011; 16(5):588–611.

2. Halkitis PN, Krause KD, Vieira DL. Mental Health, Psychosocial Challenges and Resilience in Older Adults Living with HIV. HIV and Aging. 42: Karger Publishers; 2016. p. 187–203.

3. Chibanda D, Benjamin L, Weiss HA, Abas M. Mental, neurological, and substance use disorders in people living with HIV/AIDS in low- and middle-income countries. Journal of acquired immune deficiency syndromes (1999). 2014; 67 Suppl 1:S54–67.

4. Phillips KD, Moneyham L, Thomas SP, Gunther M, Vyavaharkar M. Social context of rural women with HIV/AIDS. Issues in mental health nursing. 2011; 32(6):374–81. https://doi.org/10.3109/01612840.2011.568273 PMID: 21692576

5. Rao D, Kekwaletswe TC, Hosek S, Martinez J, Rodriguez F. Stigma and social barriers to medication adherence with urban youth living with HIV. AIDS Care. 2007; 19(1):28–33. https://doi.org/10.1080/09540120600652303 PMID: 17129855

6. Parker R, Aggleton P. HIV and AIDS-related stigma and discrimination: a conceptual framework and implications for action. Social science & medicine. 2003; 57(1):13–24.

7. Feyissa GT, Abebe L, Girma E, Woldie M. Stigma and discrimination against people living with HIV by healthcare providers, Southwest Ethiopia. BMC Public Health. 2012; 12:522. https://doi.org/10.1186/1471-2458-12-522 PMID: 22794201

8. Nyblade L, Stangl A, Weiss E, Ashburn K. Combating HIV stigma in health care settings: what works? Journal of the International AIDS Society. 2009; 12:15-. https://doi.org/10.1186/1758-2652-12-15 PMID: 19660113

9. Carr D, Kidd R., and Nyblade L. Achieving a Stigma-free Health Facility and HIV Services: Resources for Administrators. Washington, DC: Futures Group, Health Policy Project; 2015.

10. Jordan; Z, Lockwood; C, Aromataris; E, Munn Z. The updated JBI model for evidence-based healthcare.: The Joanna Briggs Institute; 2016.

11. Ciliska D, Thomas H, Buffett C. An introduction to evidence-informed public health and a compendium of critical appraisal tools for public health practice. National Collaborating Centre for Methods and Tools. 2008.

12. Francke AL, Smit MC, de Veer AJE, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: A systematic meta-review. BMC Med Inform Decis Mak. 2008; 8.

13. Shekelle P, Woolf S, Grimshaw JM, Schünemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implementation Science. 2012; 7(1):62.

14. NICE. Developing NICE guidelines: the manual: National Instutites for Health and Clinical Excellence.; 2014 [01/02.2017]. https://www.nice.org.uk/process/pmg20/chapter/introduction-and-overview.

15. Harrison M, Van Den Hoek J. CAN-IMPLEMENT: A guideline adaptation and implementation planning resource. Kingston, Ontario: Queen's University School of Nursing and Canadian Partnership Against Cancer. 2012.

16. NCCMT. Critically appraising practice guidelines: The AGREE II instrument. Hamilton, ON: McMaster University. (Updated 01 November, 2013). 2011 http://www.nccmt.ca/registry/view/eng/100.html.

17. Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak. 2005; 5.

18. Kashyap N, Dixon J, Michel G, Brandt C. GuideLine Implementability Appraisal v. 2.0 New Haven, CT: Yale University: Yale Center for Medical Informatics; 2005 http://gem.med.yale.edu/glia/doc/GLIA_v2.pdf.

19. Dalkey N, Helmer O. An experimental application of the Delphi method to the use of experts. Management science. 1963; 9(3):458–67.

20. Vernon W. The Delphi technique: A review. International Journal of Therapy & Rehabilitation. 2009; 16(2):69–76 8p.

21. Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. Semin Arthritis Rheum. 2011; 41(2):95–105. https://doi.org/10.1016/j.semarthrit.2010.12.001 PMID: 21420149

22. James D, Warren-Forward H. Research methods for formal consensus development. Nurse researcher. 2015; 22(3):35–40. https://doi.org/10.7748/nr.22.3.35.e1297 PMID: 25587865

23. von der Gracht HA. Consensus measurement in Delphi studies: Review and implications for future quality assurance. Technological Forecasting and Social Change. 2012; 79(8):1525–36.

24. Yousuf MI. Using experts' opinions through Delphi technique. Practical assessment, research & evaluation. 2007; 12(4):1–8.

25. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. Int J Clin Pharm. 2016.

26. Franklin KK, Hart JK. Idea Generation and Exploration: Benefits and Limitations of the Policy Delphi Research Method. Innovative Higher Education. 2006; 31(4):237–46.

27. Hsu C-C, Sandford BA. Minimizing non-response in the Delphi process: How to respond to non-response. Practical Assessment, Research & Evaluation. 2007; 12(17):62–78.

28. Loblaw DA, Prestrud AA, Somerfield MR, Oliver TK, Brouwers MC, Nam RK, et al. American Society of Clinical Oncology Clinical Practice Guidelines: Formal Systematic Review–Based Consensus Methodology. Journal of Clinical Oncology. 2012; 30(25):3136–40. https://doi.org/10.1200/JCO.2012.42.0489 PMID: 22778311

29. Akins RB, Tolson H, Cole BR. Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. BMC Med Res Methodol. 2005; 5:37. https://doi.org/10.1186/1471-2288-5-37 PMID: 16321161

30. Brosseau L, Wells GA, Pugh AG, Smith CA, Rahman P, Alvarez Gallardo IC, et al. Ottawa Panel evidence-based clinical practice guidelines for therapeutic exercise in the management of hip osteoarthritis. Clin Rehabil. 2015.

31. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, et al. Consensus development methods, and their use in clinical guideline development. Health Technol Assess. 1998; 2(3):i–iv, 1–88. PMID: 9561895

32. van der Linde H, Hofstad CJ, van Limbeek J, Postema K, Geertzen JH. Use of the Delphi Technique for developing national clinical guidelines for prescription of lower-limb prostheses. J Rehabil Res Dev. 2005; 42(5):693–704. PMID: 16586195

33. Thangaratinam S, Redman CW. The delphi technique. The obstetrician & gynaecologist. 2005; 7(2):120–5.

34. Green B, Jones M, Hughes D, Williams A. Applying the Delphi technique in a study of GPs' information requirements. Health & social care in the community. 1999; 7(3):198–205.

35. Williams PL, Webb C. The Delphi technique: a methodological discussion. Journal of advanced nursing. 1994; 19(1):180–6. PMID: 8138622

36. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. Information & management. 2004; 42(1):15–29.

37. Loo R. The Delphi method: a powerful tool for strategic management. Policing: An International Journal of Police Strategies & Management. 2002; 25(4):762–9.

38. Hsu C-C, Sandford BA. The Delphi technique: making sense of consensus. Practical assessment, research & evaluation. 2007; 12(10):1–8.

39. Fleuren M, Wiefferink K, Paulussen T. Determinants of innovation within health care organizations: literature review and Delphi study. International journal for quality in health care: journal of the International Society for Quality in Health Care. 2004; 16(2):107–23.

40. Goodman CM. The Delphi technique: a critique. Journal of advanced nursing. 1987; 12(6):729–34. PMID: 3320139

41. Dajani JS, Sincoff MZ, Talley WK. Stability and agreement criteria for the termination of Delphi studies. Technological forecasting and social change. 1979; 13(1):83–90.

42. Du Plessis E, Human SP. The art of the Delphi technique: highlighting its scientific merit. Health SA Gesondheid. 2007; 12(4).

43. Murry JW, Hammons JO. Delphi: A versatile methodology for conducting qualitative research. The Review of Higher Education. 1995; 18(4):423.

44. Sharkey SB, Sharples AY. An approach to consensus building using the Delphi technique: developing a learning resource in mental health. Nurse education today. 2001; 21(5):398–408. https://doi.org/10.1054/nedt.2001.0573 PMID: 11403587

45. Day J, Bobeva M. A Generic Toolkit for the Successful Management of Delphi Studies. Electronic Journal of Business Research Methods. 2005; 3(2):103–16.

46. Egon GG. ERIC/ECTJ Annual Review Paper: Criteria for Assessing the Trustworthiness of Naturalistic Inquiries. Educational Communication and Technology. 1981; 29(2):75–91.

47. Engels TC, Kennedy HP. Enhancing a Delphi study on family-focused prevention. Technological forecasting and social change. 2007; 74(4):433–51.

48. Cornick P. Nitric oxide education survey–Use of a Delphi survey to produce guidelines for training neonatal nurses to work with inhaled nitric oxide. Journal of Neonatal Nursing. 2006; 12(2):62–8.

49. Guba EG. CRITERIA FOR ASSESSING THE TRUSTWORTHINESS OF NATURALISTIC INQUIRIES. Ectj-Educational Communication and Technology Journal. 1981; 29(2):75–91.

50. Cole FL. Content analysis: process and application. Clin Nurse Spec. 1988; 2(1):53–7. PMID: 3349413

51. Elo S, Kyngas H. The qualitative content analysis process. Journal of advanced nursing. 2008; 62(1):107–15. https://doi.org/10.1111/j.1365-2648.2007.04569.x PMID: 18352969

52. Kastner M, Estey E, Bhattacharyya O. Better guidelines for better care: enhancing the implementability of clinical practice guidelines. Expert Review of Pharmacoeconomics & Outcomes Research. 2011; 11(3):315–24.

53. Chaffin WW, Talley WK. Individual stability in Delphi studies. Technological Forecasting and Social Change. 1980; 16(1):67–73.

54. Yang YN. Testing the Stability of Experts' Opinions between Successive Rounds of Delphi Studies. 2003.

55. Kalaian SA, Kasim RM. Terminating sequential Delphi survey data collection. Practical Assessment, Research & Evaluation. 2012; 17(5):1–10.

56. Zhang W, Moskowitz R, Nuki G, Abramson S, Altman R, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis and cartilage. 2008; 16(2):137–62. https://doi.org/10.1016/j.joca.2007.12.013 PMID: 18279766