Original Empirical Research

Training primary health care providers in Colombia, Mexico and Peru to increase alcohol screening: Mixed-methods process evaluation of implementation strategy

Daša Kokole[1] , **Eva Jané-Llopis[1,2,3]**, **Guillermina Natera Rey[4]**,
Natalia Bautista Aguilar[4] , **Perla Sonia Medina Aguilar[4]**,
Juliana Mejía-Trujillo[5], **Katherine Mora[5]**, **Natalia Restrepo[5]**,
Ines Bustamante[6], **Marina Piazza[6]**, **Amy O'Donnell[7]**,
Adriana Solovei[1], **Liesbeth Mercken[1,8]**, **Christiane
Sybille Schmidt[9]**, **Hugo Lopez-Pelayo[10,11,12]**, **Silvia Matrai[10]**,
Fleur Braddick[10], **Antoni Gual[10,11,12]**, **Jürgen Rehm[3,13,14,15,16]**,
Peter Anderson[1,7] **and Hein de Vries[1]**

Abstract

Background

Initial results from the SCALA study demonstrated that training primary health care providers is an effective implementation strategy to increase alcohol screening in Colombia, Mexico and Peru, but did not show evidence of superior performance for the standard compared to the shorter training arm. This paper elaborates on those outcomes by examining the relationship of training-related process evaluation indicators with the alcohol screening practice.

[1]Department of Health Promotion, CAPHRI Care and Public Health Research Institute, Maastricht University, POB 616, 6200 MD, Maastricht, The Netherlands
[2]Univ. Ramon Llull, ESADE, Avenida de Pedralbes, 60, 62, 08034 Barcelona, Spain
[3]Institute for Mental Health Policy Research, CAMH, 33 Russell Street, Toronto, ON M5S 2S1, Canada
[4]Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Calz México-Xochimilco 101, Huipulco, 14370 Ciudad de México, CDMX, Mexico
[5]Corporación Nuevos Rumbos, Calle 108 A # 4-15, Bogotá, Colombia
[6]School of Public Health and Administration, Universidad Peruana Cayetano Heredia, Ave. Honorio Delgado 430, Urb. Ingeniería, S.M.P. Lima - Perú
[7]Population Health Sciences Institute, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne NE2 4AX, UK
[8]Department of Health Psychology, Open University, Valkenburgerweg 177, 6419 AT Heerlen, the Netherlands
[9]Centre for Interdisciplinary Addiction Research (ZIS), Department of Psychiatry and Psychotherapy, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

[10]Addictions Unit, Psychiatry Dept, Hospital Clínic, Villarroel 170, 08036 Barcelona, Spain
[11]Red de Trastornos Adictivos, Instituto Carlos III, Sinesio Delgado, 4, 28029 – Madrid, Spain
[12]Institut d'Investigacions Biomèdiques August Pi Sunyer (IDIBAPS), Rosselló, 149-153, 08036 Barcelona, Spain
[13]Institute for Clinical Psychology and Psychotherapy, TU Dresden, Chemnitzer Str. 46, 01187 Dresden, Germany
[14]Dalla Lana School of Public Health, University of Toronto, 155 College Street, 6th Floor, Toronto, ON M5T 3M7, Canada
[15]Department of Psychiatry, University of Toronto, 250 College Street, 8th Floor, Toronto, ON M5T 1R8, Canada
[16]Department of International Health Projects, Institute for Leadership and Health Management, I.M. Sechenov First Moscow State Medical University, Trubetskaya str., 8, b. 2, 119992, Moscow, Russian Federation

Corresponding author:
Daša Kokole, Department of Health Promotion, CAPHRI Care and Public Health Research Institute, Maastricht University, POB 616, 6200 MD, Maastricht, The Netherlands.
Email: d.kokole@maastrichtuniversity.nl

Methods

A mix of convergent and exploratory mixed-methods design was employed. Data sources included training documentation, post-training questionnaires, observation forms, self-report forms and interviews. Available quantitative data were compared on outcome measure – providers' alcohol screening.

Results

Training reach was high: three hundred fifty-two providers (72.3% of all eligible) participated in one or more training or booster sessions. Country differences in session length reflected adaptation to previous topic knowledge and experience of the providers. Overall, 49% of attendees conducted alcohol screening in practice. A higher dose received was positively associated with screening, but there was no difference between standard and short training arms. Although the training sessions were well received by participants, satisfaction with training and perceived utility for practice were not associated with screening. Profession, but not age or gender, was associated with screening: in Colombia and Mexico, doctors and psychologists were more likely to screen (although the latter represented only a small proportion of the sample) and in Peru, only psychologists.

Conclusions

The SCALA training programme was well received by the participants and led to half of the participating providers conducting alcohol screening in their primary health care practice. The dose received and the professional role were the key factors associated with conducting the alcohol screening in practice.

Plain Language Summary: Primary health care providers can play an important role in detecting heavy drinkers among their consulting patients, and training can be an effective implementation strategy to increase alcohol screening and detection. Existing training literature predominantly focuses on evaluating trainings in high-income countries, or evaluating their effectiveness rather than implementation. As part of SCALA (Scale-up of Prevention and Management of Alcohol Use Disorders in Latin America) study, we evaluated training as implementation strategy to increase alcohol screening in primary health care in a middle-income context. Overall, 72.3% of eligible providers attended the training and 49% of training attendees conducted alcohol screening in practice after attending the training. Our process evaluation suggests that simple intervention with sufficient time to practice, adapted to limited provider availability, is optimal to balance training feasibility and effectiveness; that booster sessions are especially important in context with lower organizational or structural support; and that ongoing training refinement during the implementation period is necessary.
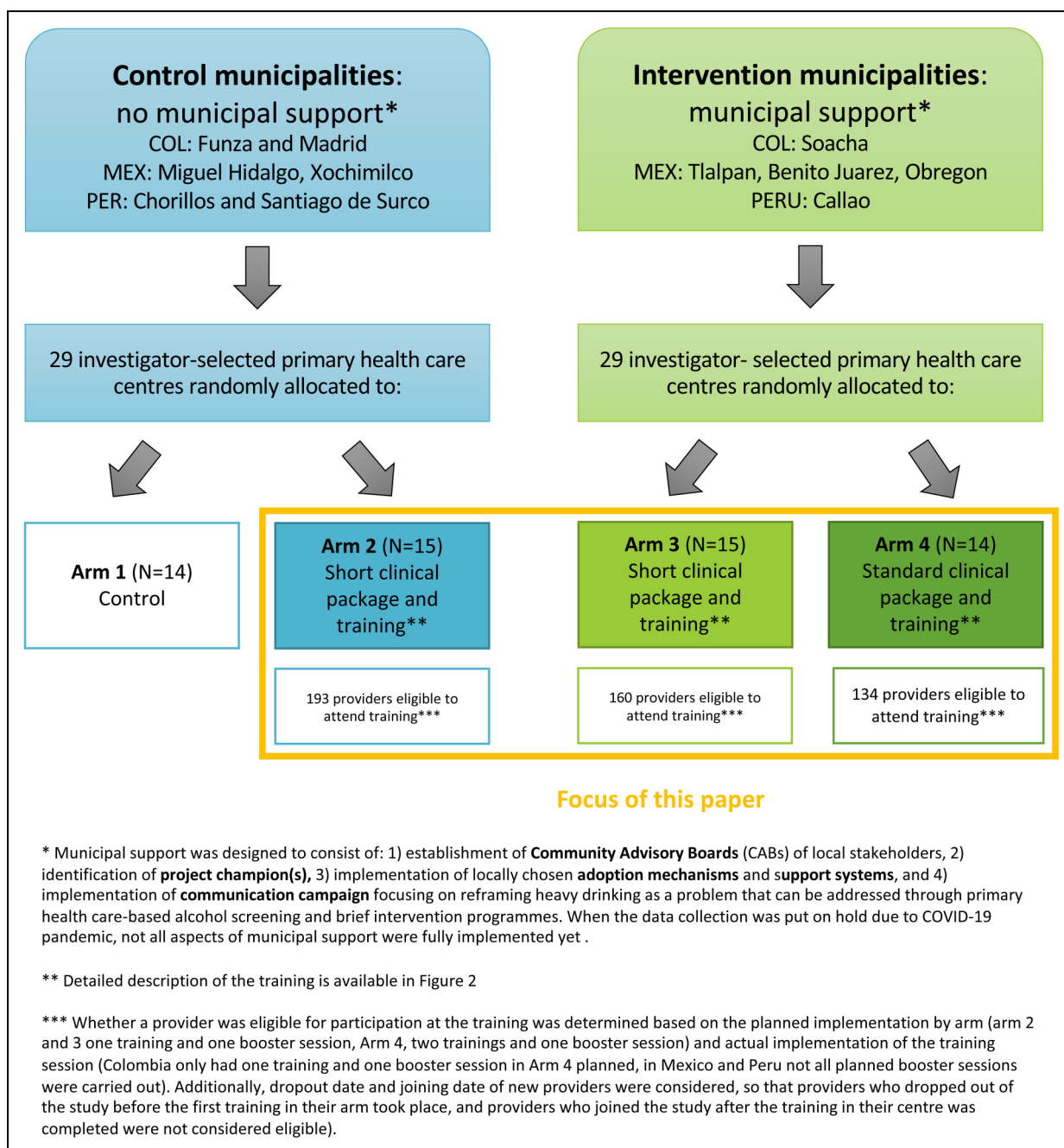
Introduction

Primary health care (PHC) providers are a key group that can encourage adults to adopt a healthier lifestyle, as they have regular access to large portions of the population through routine consultations. One of the key components of a healthy lifestyle is reduction of alcohol use, alcohol being the ninth-largest risk factor for morbidity globally, and fourth in Latin America (GBD 2019 Risk Factors Collaborators, 2020). Screening for patients' risky alcohol use during a routine check-up in PHC and providing them with advice on cutting down if necessary has a large body of evidence supporting its effectiveness in the reduction of alcohol use (Kaner et al., 2018; O'Donnell, Anderson, et al., 2014). Despite its effectiveness and simplicity, it is often sub-optimally implemented in practice (O'Donnell, Wallace, et al., 2014) and a large body of research deals with how to improve the implementation to achieve better public health outcomes (Anderson et al., 2016; Keurhorst et al., 2016; Heather, 2006).

An implementation strategy consistently shown to be effective in improving the implementation of healthcare interventions in general (Rowe et al., 2018), and alcohol screening in particular (Anderson et al., 2014) is training of the PHC providers. Conversely, lack of appropriate training has emerged as a major barrier to practice; in a recent systematic review of factors influencing PHC providers' implementation of alcohol screening and brief intervention in primary care practices, lack of training was the most commonly emerging theme among the cited barriers, closely followed by the alcohol-related knowledge and the belief in one's own ability to deliver the intervention - both of which can be also targeted via training (Rosário et al., 2021). However, the majority of the existing alcohol screening research (both in general and training-specific) comes from high-income countries (Rosário et al., 2021). While alcohol screening programmes have been implemented and evaluated in Latin America (Ronzani et al., 2019; Shannon et al., 2021), evaluation of providers' training research remains scarce (e.g. (Furtado et al., 2008). Training evaluation in other healthcare-related fields focuses on effectiveness rather than the implementation (Smith et al., 2020; Spagnolo et al.,

Figure 1. SCALA study design.

* Municipal support was designed to consist of: 1) establishment of **Community Advisory Boards** (CABs) of local stakeholders, 2) identification of **project champion(s)**, 3) implementation of locally chosen **adoption mechanisms** and s**upport systems**, and 4) implementation of **communication campaign** focusing on reframing heavy drinking as a problem that can be addressed through primary health care-based alcohol screening and brief intervention programmes. When the data collection was put on hold due to COVID-19 pandemic, not all aspects of municipal support were fully implemented yet .

** Detailed description of the training is available in Figure 2

*** Whether a provider was eligible for participation at the training was determined based on the planned implementation by arm (arm 2 and 3 one training and one booster session, Arm 4, two trainings and one booster session) and actual implementation of the training session (Colombia only had one training and one booster session in Arm 4 planned, in Mexico and Peru not all planned booster sessions were carried out). Additionally, dropout date and joining date of new providers were considered, so that providers who dropped out of the study before the first training in their arm took place, and providers who joined the study after the training in their centre was completed were not considered eligible).

2020; Stokholm Bækgaard et al., 2021; Stoltenberg et al., 2020; Suriyawongpaisal et al., 2020). This paper addresses this gap in the literature and presents the findings of an in-depth process evaluation of using a training package as an implementation strategy to increase alcohol screening by the PHC providers in a middle-income context.

SCALA (Scale-up of Prevention and Management of Alcohol Use Disorders in Latin America) is a quasi-experimental study conducted in three middle-income Latin American countries (Colombia, Mexico and Peru),

testing whether training PHC providers and providing municipal support (a range of adoption mechanisms and support systems) could support improved implementation of PHC-based screening, advice and treatment for heavy drinking and comorbid depression (Jane-LLopis et al., 2020). In addition to screening for risky alcohol consumption of all patients, and providing advice on cutting down if necessary, SCALA clinical care pathway requires providers to check for depressive symptoms in the heavy drinking patients, as heavy drinking is often comorbid with depression (Bellos et al., 2016), and

Table 1. Description of the setting characteristics in Colombia, Mexico and Peru.

	Colombia	Mexico	Peru
Setting description			
Main country demographics	• Population: 48 258 494 (2018 data) • 51.2% female • 75.5% living in urban areas • Age distribution: 24.0% under 15, 67% 15–64, 8.8% 65 +[a]	• Population 119 938 473 (2015 data) • 51.4% female • 76.8% living in urban areas • Age distribution (2010 data): 29.3% under 15, 64.4% 15–64, 6.3% 65 +[b]	• Population 31 237 385 (2017 data) • 50.5% female • 81.9% living in urban areas • Age distribution: 26.5% under 15, 65.3% 15–64, 8.2% 65 +[c]
GDP per capita (2019)[d]	6508.1 USD	10118.2 USD	7046.8 USD
Income level (World bank)[e]	Upper-middle income	Upper-middle income	Upper-middle income
Participating municipalities	Intervention (Arm 3 and 4): Soacha (population: 93,154; located in the metropolitan area of Bogota, part of the department of Cundinamarca).[a] Control (Arm 1 and 2); Madrid (93,154), Funza (pop: 112,254), both located in Western Savanna Province and part of the department of Cundinamarca, 25 km outside Bogota.[a]	Intervention: Tlalpan (650,567)*; Benito Juárez (385,439), Álvaro Obregón (727,034); all one of 16 municipalities of Mexico City.[b] Control: Miguel Hidalgo (372,889), Xochimilco (415,007), both one of 16 municipalities of Mexico City.[b] *two of PHCUs from this municipality are in the control arm	Intervention: Callao (pop: 451,260): Provincial capital and one of the seven districts in Callao province, part of Callao region. Located in the West area of Lima, and borders the Pacific Ocean.[c] Control: Chorrillos (314,241) and Santiago de Surco (329,152); both one of the 43 districts of Lima province, located in the Lima region, bordering each other.[c]
Existing alcohol SBI practice and guidelines	The alcohol SBI recommendations are included as part of clinical practice guidelines that focus on detection and treatment of alcohol abuse and dependence on primary, secondary and tertiary care level[f] but there are no official standards. Some of the providers are familiar with the screening instruments.	Official standards establish the obligatory procedures and criteria for mandatory prevention, treatment and control of addictions, which include asking questions on alcohol use[g] and including this information in the patient's history[h] specifically in primary health care context.	No explicit guidelines, recommendation for providers to include alcohol screening is implicitly included in general recommendation to perform mental health related screening (alcohol use disorder being considered as one of subcategories)[i]
Organisational context in the participating PHCCs[j]	In the intervention arm, the leadership was very supportive of the project, and there was leadership directive for providers' participation and assigned time to attend the trainings.	The organisational context depended on the centres, with varying levels of leadership support. In all of the participating centres there was existing screening practice due to standards described above, and providers were familiar with the screening instruments and often also have experience with its application.	There was no consistent leadership directive in the centres. None of the participating centres had an existing screening practice and providers were generally not familiar with screening instruments.
Provider recruitment details	In all arms, the providers were chosen for participation by their superiors.	The recruitment varied by centre, with some providers being selected for participation and some providers volunteering to participate.	The providers had to volunteer to join the project.

[a]DANE (2018). Censo nacional de población y vivienda. Proyecciones de población. Available from: https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion [accessed 23.9.2020]. [b] INEGI (n.d.). Banco de indicadores, 2015. Available from https://www.inegi.org.mx/app/indicadores/?t = 0070&ag = 09014###D00700060 [accessed 23.9.2020] [c] INEI (2017). Censos nacionales 2017: XII Censo de Población, VII de Vivienda y III de Comunidades Indígenas. Sistema de Consulta de Base de Datos. Available from: http//censos2017.inei.gob.pe/redatam/ [accessed 23.9.2020] (data from 2017). [d] IMF (2019). World Economic Outlook: https://www.imf.org/en/Publications/SPROLLS/world-economic-outlook-databases. [e] World bank (n.d.). World Bank Country and lending groups: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519. [f] Ministerio de Salud y Protección Social. Guía de práctica clínica para la detección temprana, diagnóstico y tratamiento de la fase aguda de intoxicación de pacientes con abuso o dependencia del alcohol - 2013 Guía No. 23 [Internet]. 2013. Available from: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/INEC/IETS/GPC_Completa_OH.pdf. [g] Norma Oficial Mexicana NOM-028-SSA2-2009 para la prevención, tratamiento y control de las adicciones [Internet]. 2009. Available from: http://www.cndh.org.mx/DocTR/2016/JUR/A70/01/JUR-20170331-NOR26.pdf. [h] Norma Oficial Mexicana NOM-004-SSA3-2012 del expediente clínico [Internet]. 2012. Available from: https://www.conadic.salud.gob.mx/pdfs/norma_oficial_nom.pdf. [i] Ministerio de Salud Perú. Plan nacional de fortalecimiento de servicios de salud mental comunitaria 2018-2021 [Internet]. 2018. Available from: http://bvs.minsa.gob.pe/local/MINSA/4422.pdf. [j]Information provided by local research partners based on field visits.

associated with worsening of depression, including increased suicide risk and impaired social functioning (Boden & Fergusson, 2011). A summary of the study design and the included intervention components by arm is presented in Figure 1, and further detailed description is available in the protocol paper (Jane-LLopis et al., 2020). The results of the outcome evaluation at the primary health care center (PHCC) level have been published elsewhere and the findings suggest that the providers' training significantly increased the proportion of alcohol screening in adult patients (although the standard training and clinical package was not superior to shorter version) (Anderson et al., 2021), as well as the depression screening rates (O'Donnell et al., 2021). At the time of evaluation, municipal support (as described in Figure 1), was not found to have a significant impact (Anderson et al., 2021), which was likely due to its incomplete implementation, as the implementation had to be put on hold at the 5-month mark due to COVID-19 restrictions.

In this paper, we present the findings from the process evaluation of SCALA training as the implementation strategy already demonstrated to be effective to increase alcohol screening (Anderson et al., 2021). We used the UK Medical Research Council (MRC)'s process evaluation guidance (Moore et al., 2014, 2015) to develop the process evaluation protocol (Jane-LLopis et al., 2020) and guide the aims, focusing on the key issues regarding implementation, mechanisms of impact, and context along with their relationship to the outcome. The main research questions addressed in this paper include:

- What was implemented in terms of training dose, reach, adaptation and fidelity?
- How did the participants respond to the training, and were there any unintended consequences of the training implementation?
- How were the implementation factors, participant response and provider demographics associated with the main study outcome (alcohol screening in practice)?

Methods

Design and Setting of the Study

We used the StaRI checklist (Pinnock et al., 2017) to report on the study. The presented data have been collected as part of the broader process evaluation (Jane-LLopis et al., 2020) to support an in-depth understanding of study's primary results (Anderson et al., 2021). We employed mixture of the convergent and exploratory mixed-methods design (Creswell & Plano Clark, 2011), whereby the qualitative and quantitative data were collected in the same period, analyzed separately, but then complemented with additional qualitative data and interpreted together. Some relevant characteristics of the implementation setting are presented in Table 1.

SCALA Training Curriculum

Figure 2 describes the SCALA training curriculum. Two versions of the training package were developed, a short and standard version, to be tested in different arms. Both were designed to be flexible and easy to adapt to the country and local context. The training package consisted of four products: the training manual; the training course presentations; the training videos, and the Train New Trainers sessions. The key differences between the short and standard training package were different care pathways (short vs. standard, as described in the study protocol (Jane-LLopis et al., 2020)), and a different set of videos. Both training sessions focused on alcohol screening and advice, and the standard training additionally emphasized dealing with co-morbid depressive symptoms, and alcohol referral and treatment options. The core of the training sessions was based on two main components: videos showing PHC providers delivering the protocol in practice and role-plays using the developed materials. The theoretical background for this approach comes from social cognitive theory (Bandura, 1977), where both vicarious learning (through modelling) and enactive mastery (practicing the skills yourself) are key components of increasing self-efficacy. In practice, this approach has been used in previous similar projects (e.g. ODHIN, PHEPA) with demonstrated effectiveness (Anderson et al., 2016).

Before training implementation, Train New Trainers course took place in Bogota, Colombia in May 2018, attended by the local professionals (doctors, psychologists, addiction specialists) from the three countries (future trainers, N = 16). The training was conducted by an addiction specialist with several years of experience in implementing brief interventions and training delivery. A detailed description of the full training package and the process of its development is available in Supplementary file 1, along with the links to all the training materials (SCALA, 2021).

Participants

PHC providers of any professional role from the participating centers were eligible for participation in the SCALA study upon signing informed consent. Information on provider recruitment is described in Table 1. In this paper, we included the providers from 44 PHCCs in Arms 2, 3 and 4 who were eligible to be trained either in the baseline period, or through booster sessions taking place in the first five months of the implementation period of the study. Some centers were in their 6th or 7th month of implementation when data collection was halted due to the start of the COVID-19 pandemic and were still recruiting new providers at that point, but we used the 5-month mark to allow for better alignment with the data presented in the main outcome paper (Anderson et al., 2021). Providers from the 14 PHCCs in Arm 1 (control group) who did

Figure 2. SCALA training curriculum.

	Short training (Arm 2 and 3)	Standard training (Arm 4)
	Suggested training length: 1 x 2h training session (T1) 1 x 1h booster session (B1)	Suggested training length: 2 x 2h training sessions (T1 and T2)- 4h total 2 x 1hour (or 1x2 hours) booster sessions (B1)
T1 (all arms)		
Unit 1: General concepts + Attitudes to alcohol (30 min)	• Welcome and warmup • Quiz on the impact of alcohol use • Discussion on attitudes toward alcohol use	
Unit 2: Screening for alcohol problems and comorbid depression (50 min)	• Explanation of SCALA screening criteria and care pathway • Presentation of AUDIT-C and PHQ-2 as screening instruments • Showing two videos on screening • Role-playing in pairs with predefined patient roles	
Unit 3: Brief intervention on alcohol (45 min)	• Presentation of brief intervention steps and core skills needed • Showing video on brief intervention • Role-playing in pairs with predefined patient roles	
		T2 (Arm 4 only)
Unit 4: Recap + Advice/information for co-morbid depressive symptoms (40 min)		• Presentation of strategy for co-morbid depressive symptoms • Showing video on dealing with co-morbid depressive symptoms • Role-playing in pairs with predefined patient roles
Unit 5: Referrals for alcohol and depression (40 min)		• Summary presentation of services for alcohol and depressive symptoms (tailored to the local context), followed by discussion. • Show video on referral • Role-playing in pairs with predefined patient roles
Unit 6: Treatment when the referral is not possible + wrap up (30 min)		• Presentation on options for professionals when referral services are not available or patients are not willing
B1 (all arms)		
Booster session (60 min)	• Summary of key concepts and care pathway process • Troubleshooting based on providers' experience	

not receive training, as well as providers who attended the training without signing the informed consent to participate in the study, were not included in the data collection and analysis.

Measurements

The selection of constructs to inform the training process evaluation was based on the UK MRC process evaluation

Table 2. Key measured constructs for the process evaluation of the training.

Construct measured	Definition (MRC)[a]	Operationalization	Data source used	Data integration
Implementation				
Dose	Quantity of what has been delivered in practice (how much intervention is delivered)	Actual amount and length of the training per country and arm (dose delivered) Number of hours and sessions provider participated in the training, per country and arm (dose received)	Training documentation, attendance lists	Quantitative data only, information on dose delivered combined with individual attendance information to calculate dose received
Reach	The extent to which the target audience comes into contact with the intervention	Number and % of all providers recruited for the study and eligible for the training that participated in the trainings[b] calculated for each training session separately and overall, per country and per arm. Representativeness of the reached population assessed through a comparison of demographic characteristics (age, gender, professional role) between the reached population and non-reached population.	Training documentation, consent form information	Quantitative data on reach complemented with qualitative on reasons for provider non-participation (as obtained through interviews)
Adaptation	Alterations made to intervention to better fit the context	Description of parts of the training that were adapted	Observation forms, trainer self-report forms	Results triangulated from both data sources and complemented with qualitative data on reasons for changes
Fidelity	Quality of what is delivered or consistency of what is implemented with the planned interventions	Delivery of training's active ingredients (videos and role-plays): complete, partial or none.	Observation forms, trainer self-report forms	Results triangulated from both data sources and complemented with qualitative data on reasons for changes
Mechanisms of impact				
Participant response	How do participants interact with the intervention	Satisfaction with the training Perceived utility of the training Suggestions for improvement	Post-training questionnaire, interviews	Quantitative data complemented with qualitative data on provider's impressions and suggestions for improvements and qualitative data from trainer interviews
Unintended consequences	Unanticipated pathways and consequences	Emerging side-effects of delivering trainings in the PHCC	Interviews	Qualitative data only
Context				
Demographics	Factors external to the intervention which impede or strengthen its implementation or effects	Country, age, gender, professional role of providers	Consent forms	Quantitative data only
Outcome				
Outcome behaviour	Participating providers doing alcohol screening in practice at least once		Tally sheets	Quantitative data only, integrated with dose, participant response and demographic information

[a]In this paper, term intervention (as used in the MRC definition) refers to the training package. [b] Whether a provider was eligible for participation at the training was determined based on the planned implementation by arm (arm 2 and 3 one training and one booster session, Arm 4, two trainings and one booster session) and actual implementation of the training session (Colombia only had one training and one booster session in Arm 4 planned, in Mexico and Peru not all planned booster sessions were carried out). Additionally, dropout date and joining date of new providers were considered, so that providers who dropped out of the study before the first training in their arm took place, and providers who joined the study after the training in their centre was completed were not considered eligible).
Satisfaction with the training.
Perceived utility of the training.
Suggestions for improvement.

guidance (Moore et al., 2015). Table 2 presents an overview of the key measured constructs, the data sources used for their assessment, and information on data integration. We included also an outcome indicator – using SCALA clinical package in practice at least once. As the SCALA care pathway was designed for the providers to screen for alcohol consumption first and only use depression screening if the patient was identified as drinking at a risky level, alcohol screening was used as a proxy measure to indicate the use of the SCALA clinical package.

Data Sources and Collection

Details on the data sources are presented below along with data collection procedure description.

Training Documentation

Training logbooks were completed by the local research partners throughout the implementation period and contained information on the delivered training (date, time and training location) and participating providers (based on the information from the signed attendance lists). This allowed us to assess the dose and reach of the training; the latter also in combination with demographic data gathered during study recruitment.

Post-training Questionnaire

The questionnaire assessed participant response to the training. Participants answered a set of questions on a 5-point Likert scale (1-Very negative to 5-very positive for satisfaction and 1-Not very useful to 5-Very useful for perceived utility) and additionally had space to provide open-ended answers. Providers completed the pen-and-paper questionnaires at the end of the training session in the period

between August 2019 and March 2020. Questionnaires contained the predetermined provider ID to guarantee anonymity and individual traceability.

Observations

The training sessions were observed by previously trained local research team members. In Colombia and Peru, all the sessions were observed. In Mexico, one session per arm was observed due to resource limitations. Researchers used a structured observation form containing both quantitative indicators (e.g. checklists to mark the delivery of listed active ingredients to assess fidelity), and there was room for qualitative observations (e.g. providing an additional explanation in case of non- or partial execution of activity).

Self-report Forms

The trainers completed the self-report form after each delivered training, providing information on which components were delivered and whether they adapted the training, along with explanations. Fidelity and adaptation questions in the observation and self-reports form were the same to allow for data triangulation.

Interviews

The interview topic guide was tailored to the country to complement the information obtained from other sources based on previous familiarization with data from other sources by the interviewer. The interviews were conducted after data from other sources (both qualitative and qualitative) was already partially analyzed. The initial interview guides were developed as part of process evaluation protocol development, and were adapted to reflect any additional issues that emerged during the data familiarization and preliminary analysis phase. In total, three group interviews (one per country)

Table 3. Reach and dose by country.

	Colombia	Mexico	Peru
Reach	67 (89.3% eligible) providers attended at least one session	139 (65.0% eligible) providers attended at least one session	146 (73.3% eligible) providers attending at least one session
Dose delivered	Total: 16 sessions (7 T1 + 9 B)	Total: 26 sessions (18 T1,T2 + 8 B)	Total: 33 sessions (20 T1,T2 + 13 B)
	3.5 h (1.5 T1 + 2 B) - Arm 2 and 3	3 h (2 T1 + 1 B) - Arm 2 and 3	4 h (2 T1 + 2 B) - Arm 2 and 3
	4 h (2 T1 + 2 B) - Arm 4	5 h (2 T1 + 1 T2 + 2 B) - Arm 4	6 h (2 T1 + 2 T2 + 2 B) - Arm 4
Dose delivered - COVID-19 impact	All training and booster sessions delivered before the start of the COVID-19 pandemic.	Six booster sessions not delivered in part due to the COVID-19 pandemic (sessions were hard to schedule because of lower priority in the centres, and then had to be further postponed to COVID-19)	Two booster sessions not delivered in part due to the COVID-19 pandemic (sessions were first postponed due to scheduling issues and lack of trainers' time and then had to be cancelled due to COVID).
Dose received	On average, the providers participated:	On average, the providers participated:	On average, the providers participated:
	2.3h Arm 2	1.9h Arm 2,	2.6h Arm 2,
	2.4h Arm 3	2.1h Arm 3,	2.6h Arm 3,
	3.1h Arm 4	2.8h Arm 4	4.2h Arm 4

Note. T1 – first training session (all arms), T2 – second training session (only Arm 4), B – booster session.

were conducted with a total of nine participants (two in Colombia, two in Peru and five in Mexico). All participants were either trainers (N = 7) or training organizers (N = 2) in their countries.

Tally Sheets (As Outcome Data)

During the implementation period, providers completed a tally sheet each time they conducted a screening. Those were collected from the PHCCs on monthly basis by the local researchers.

The majority of the data was collected between August 2019 and March 2020, with exception of the interviews, conducted between November and December 2020. The first part of the data was collected by the local research teams (one in each of the countries), and all the evaluation materials were transferred electronically to the evaluation coordinator using 256-bit 'Advanced Encryption Standard' (AES). The online interviews were conducted by the process evaluation leader, audio-recorded with the consent of all participants, transcribed in Spanish and translated to English. All recordings were destroyed after transcription.

Data Analysis

Quantitative data were analyzed with SPSS 25 (IBM Corp., 2017). Frequencies and descriptive statistics (mean, standard deviation) were calculated and group comparisons were made using Mann-Whitney U and Chi-square tests - overall, and by country or arm where applicable. Qualitative data was analyzed based on a combination of

inductive and deductive coding using Atlas.ti 8.4 (Scientific Software Development GmbH, 2019). The main framework for deductive coding of qualitative parts of post-training questionnaires, observation forms and self-report forms were the categories based on the MRC guide as presented in Table 3, and within those categories, themes were coded inductively. Any discrepancies between information in observation and self-report forms were resolved through discussion with the local research teams. To code the interviews, independent double coding was conducted based on the framework by two researchers (DK and AS), PhD candidates with a background in health promotion/health communication and implementation science, followed by a coding comparison and summary of the main emerging themes. The lead author integrated the quantitative and qualitative data by the framework category with the purpose of complementarity (elaboration and clarification of the results obtained through one method with results from another) (Greene et al., 1989). Table 2 gives a more detailed indication on how the data were integrated at the point of analysis.

Ethics

The study has been reviewed and approved by the research ethics board at the TU Dresden, Germany (registration number: 'EK 90032018'), and by the ethics boards in Colombia, Mexico, and Peru. All participating providers signed informed consent upon study recruitment.

Figure 3. Summary of key findings of SCALA training process evaluation.
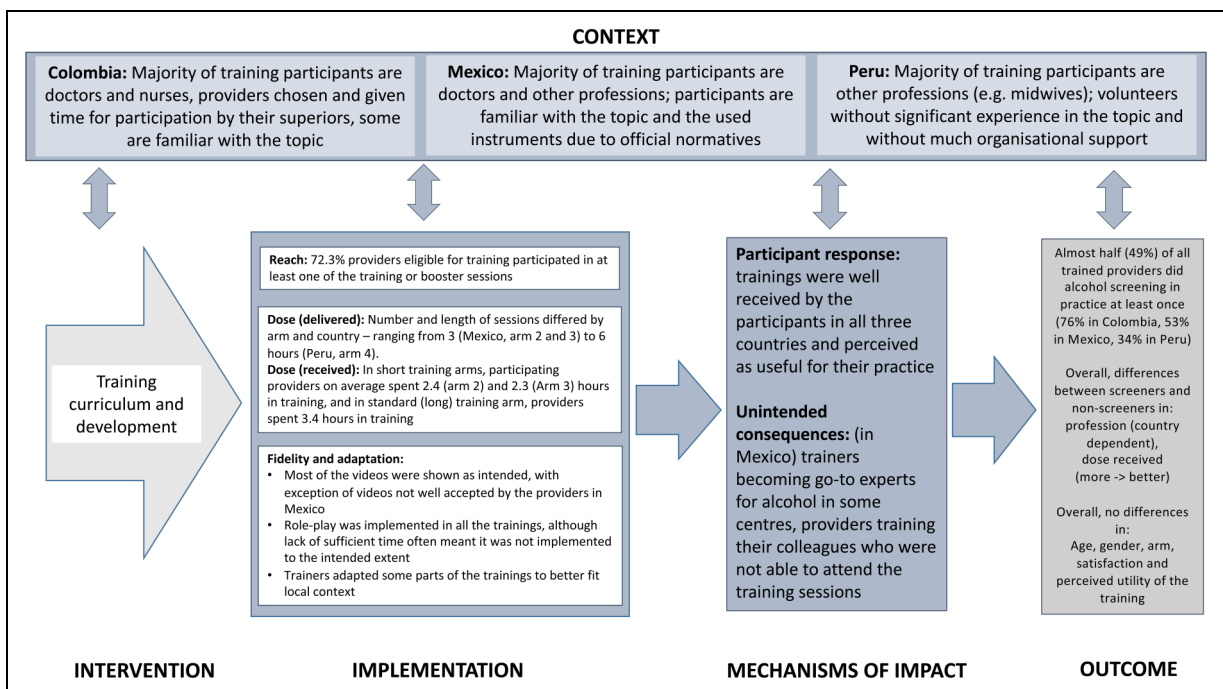
Results

The results of the key process indicators and their relationship with the outcome are presented below. Due to the large amount of collected and analyzed data, only the key tables are included in the results section, remaining tables are available as supplementary material (Supplementary file 2). A summary of the key findings is presented in Figure 3.

The training sessions took place between August and November 2019, and booster sessions took place between January and March 2020. In total, 45 training sessions and 30 booster sessions were delivered by twelve trainers (three from Colombia, five from Mexico, and four from Peru). Most of the activities were carried out before the restrictions due to the COVID-19 pandemic (Table 3).

Reach and Dose

First, we calculated reach (how many providers attended the training), dose delivered (how much training was offered) and dose received (how much training the providers attended) based on the information from the training documentation and attendance lists. A summary of the main reach and dose indicators is presented in Table 3, with complete information on reach and dose by country and arm available in Supplementary file 2, Table S2. Overall, almost three quarters (72.3%) of all eligible providers attended at least one of the delivered training sessions, with the highest percentage in Arm 4 (76.9%), followed by Arm 2 (74.1%) and Arm 3 (66.3%). In Arm 2 and 3, providers in all three countries had the opportunity to attend two sessions (one main and one booster), and in Arm 4, three sessions (with exception of Colombia, where two sessions were provided). Length of sessions differed by arm and country, ranging from three (Mexico, arms 2 and 3) to six hours (Peru, arm 4). The average number of hours attended across countries was highest in standard training in Arm 4 (3.4), and was comparable between short training arms (Arm 2 and 3); 2.4 and 2.3 h respectively. On average across arms, providers in Colombia spent 2.7 h in training, in Mexico 2.2 h and in Peru, 3.1 h.

Comparing the attending and non-attending providers on demographic characteristics (Supplementary file 2, Table S3) showed no overall difference by age or professional role, but a higher percentage of eligible women attending compared to eligible men (74.8% vs. 65.1%, χ^2 = 4.40, p = 0.036). In Mexico, we found significant difference regarding professional role (χ^2 = 8.24, p = 0.041), as all the eligible psychologists, but only over half of the nurses attended the training.

As part of the interviews, we asked the local training teams about what they perceived as reasons for some providers not attending the training. Respondents indicated that lack of attendance did not always mean a lack of interest in the project; alternative reasons for non-attendance included conflicting work obligations, or not being present at work on that day (e.g. some providers only worked weekends). Additionally, in Peru, there was a three-month gap between recruitment and training, so a possible reason suggested by the trainers was that some providers forgot they registered. All providers who were allocated to training arms in SCALA were sent training presentation slides and materials however, and in Mexico, some received additional guidance from colleagues who attended the training (see also Unintended consequences).

Fidelity and Adaptation

Table 4 presents the extent to which the main elements of the training were delivered, including the *explanation of basic concepts, videos* and *role-plays,* based on the data obtained through observation and self-report forms. Overall, the *explanation of the basic concepts* was delivered fully. *Videos* were to large extent shown as intended in Colombia and Peru, except in a few cases where trainers ran out of time. In Mexico, after showing videos at the initial 8 training sessions, comments from providers suggested that the videos did not sufficiently reflect the organizational context, therefore the training team decided to replace the videos with hypothetical cases suggested by the providers themselves.

The *role-plays* were always delivered, but were commonly shortened due to lack of time, meaning that the participating providers only practiced part of the activities as a health professional, and were experiencing the remaining activities in the patient's role. In all three countries, a lack of time for role play was evident in the short training session. Besides the adaptation of role-plays and video demonstration, other components adapted by the trainers to better fit the local context were the slide deck (to reflect the changed sequences or to reduce the amount of information on one slide), as well as the introduction activities (knowledge quiz and discussion on attitudes towards alcohol).

Participant Response

The participants completed the evaluation questionnaire at the end of the training (response rate: Training 1 (T1) 95%, Training 2 (T2), 83%, Booster 1 (B1) 77%). All the participants highly rated their overall experience with the course for each of the training or booster sessions (above 4 on a 5-point scale). There was no difference between countries in overall experience with the course in either Training 1 or the booster session. In Training 2, providers in Peru had higher ratings of the overall experience with the course compared to providers from Mexico (H(1) = 7.28, p = 0.007), although both ratings were high. In Colombia, the providers were on average slightly less satisfied with the location and venue in the T1 than in Mexico and Peru ($H_{location}(2)$ = 15.97, p < 0.001; $H_{venue}(2)$ =

22.87, p < 0.001). No other major differences between the ratings of the course were found. Full post-training questionnaire results by country and arm are available in Supplementary file 2, Tables S4-S7.

In the questionnaire, the providers could also leave open-ended comments about the training, and their answers are summarized in Supplementary file 2, Table S8. Several participants expressed that they found role-plays helpful and would appreciate more practice and examples, including personalized feedback. Overall, wishing to have more time available for training was a commonly occurring comment (except for the 2nd session in Mexico). Some providers also suggested videos could better reflect the local reality. Concerning logistics, some Colombian providers noted that training location training should be closer to their workplace and the training venue could be more comfortable. Some Peruvian providers wished for more sign-up time slots, as they are working on different schedules, and training reminders. Another theme among Peruvian providers was the importance of social support– both appreciating meeting other providers with similar interests, and wishing more providers would join the training. The interviewed trainers also corroborated the providers' reports on the importance of the opportunity to practice the skills through role-plays, and emphasized the impact of familiarity with the used instruments – less familiarity was indirectly associated with less time to practice due to more time necessary for explanation.

Unintended Consequences

No negative unintended consequences were detected in any of the countries, but two positive unintended consequences emerged in Mexico. Interviews suggested that the participating trainers established a good relationship with some PHCCs through their work with SCALA, which led to continued collaboration also beyond the scope of the training and established them as 'go-to' local experts on the topic of alcohol (for example, resulting in invitations to speak at relevant events). Additionally, in some Mexican PHCCs, it emerged that the liaising providers (contact persons, who were most closely engaged with SCALA activities) engaged themselves to provide additional explanation and training to their colleagues who were not able to attend the training. Thus, also those providers received information and training from their attending colleagues, broadening the reach of the training.

Relationship of Process Evaluation Variables with the Outcome

To assess the relationship of context (demographic factors), implementation and mechanisms of impact (participant response) with the outcome, we considered the sample of providers attending minimum one training and participating in minimum one of the five implementation

period months (N = 352). We compared the providers screening in practice at least once ("screeners", N = 173) with providers not doing any alcohol screening ("non-screeners", N = 179) on characteristics reflecting the broad categories of the MRC framework, using appropriate univariate statistic test (Chi Square or Mann-Whitney U). Results are summarized in Figure 4, with complete analysis available in Tables S9-S11 in Supplementary file 2.

Comparison of the two groups by demographic characteristics showed no significant difference in age (although in all three countries the screeners were on average slightly older than non-screeners). Significant gender differences appeared only in Mexico, where 58.9% of women and 38.6% of men among the training attendees screened (χ^2 = 4.96, p = 0.026). There was also a significant difference by profession in Colombia and Peru; in Colombia, both doctors and psychologists were more likely to screen after attending the training (χ^2 = 14.53, p = 0.002), although it should be noted that in absolute terms doctors represented the largest part of training attendees. In Peru, the psychologists among the training attendees were more likely to screen than any other profession (χ^2 = 9.64, p < 0.001).

When comparing the screeners and non-screeners on implementation factors, there was a clear trend in the total sample of screeners spending more time in training, both in terms of hours and sessions participated (2.82 vs. 2.52 h of training participation, p = 0.003; and 1.58 vs. 1.36 training sessions (p = 0.001). There was no difference by arm, meaning the providers receiving the standard training were not more likely to screen than providers in the short training arms. Comparison by the participant response showed no difference between screeners and non-screeners based on their satisfaction and perceived utility of the training (except for overall satisfaction with the training in Peru).

Additionally, we also looked at alcohol screening among the non-reached providers (providers eligible for, but not present at any of the trainings). Overall, 18% of providers not present at any of the training sessions still screened in practice, most of them from Mexico.

Discussion

This paper considered the process indicators related to training primary health care providers in Colombia, Mexico and Peru to deliver alcohol management and depression programme, as well as the relationships of those indicators with the primary outcome behavior, alcohol screening. To our knowledge, this is one of the first papers examining training as an implementation strategy through the process evaluation lens while including a multi-site comparison in the middle-income context.

Process Indicators Considerations

Reach of the programme was high, with 72.3% of the eligible providers attending at least one of the offered trained

Table 4. Overview of training fidelity and adaptations by country.

	Colombia	Mexico	Peru
Fidelity - short training			
Explanation of basic concepts: • Explanation screening criteria • Present AUDIT-C and PHQ-2 • Present steps of alcohol brief intervention • Introduction of core skills	Completed fully	Completed fully	Completed fully
Showing videos • screening video (alcohol screening – negative) • screening video (alcohol screening – positive; depression screening – negative) • brief intervention video (brief intervention for alcohol)	Completed fully	Completed partially – based on feedback from first few 8 trainings, screening videos were not shown in remaining training sessions (feedback from the participants was is that it did not reflect the Mexican context)	Completed fully – with exception of one case where brief intervention video was not shown due to lack of time
Performing role plays • Screening role play in pairs, with exchange of roles • Brief intervention role play in pairs, with exchange of roles	Completed partially – due to lack of time both role-plays had to be merged in most training sessions	Completed partially – both role plays were done, but no exchange of roles	Completed partially – in some trainings, role plays had to be merged due to lack of time
Fidelity - standard training			
Explanation of basic concepts Session 1: • Explanation of screening criteria • Presentation AUDIT/ AUDIT-C and PHQ-2/9 • Presentation steps of alcohol brief intervention • Introduction of core skillsSession 2 • Explanation of strategy for the management of co-morbid depressive symptoms • Presentation of the summary of services for the treatment of depressive symptoms and problematic alcohol use. • Presentation of treatment and follow-up options when referral is not possible	Completed fully	Completed fully Note: In Mexican centres the referral pathways are well defined, so the trainer did not have to spend much time on it	Completed fully
Showing videos Session 1: • alcohol screening – negative • alcohol screening – positive; depression screening – negative • screening and brief intervention – alcohol and depression positive • brief intervention for alcoholSession 2: • brief intervention for alcohol and depression	Session 1: Completed fully Session 2: Completed fully – exception one training where one video could not be shown due to lack of time	Session 1: Completed fully Session 2: Completed partially: based on feedback from first few 8 trainings, screening videos were not shown in remaining training sessions	Session 1: Completed fully Session 2: Completed fully

(Continued)

Table 4. (Continued)

	Colombia	Mexico	Peru
• referral for alcohol problems and co-morbid depressive symptoms			
• patient not accepting referral			
Performing role plays Session 1: • Screening role play in pairs, with exchange of roles • Brief intervention role play in pairs, with exchange of roles Session 2: • Brief intervention for alcohol and depression role play in pairs, with exchange of roles • Referral role play in pairs, with exchange of roles	Session 1 and 2: completed partially - All role plays were merged, focus was on the first one. Role-plays from session 2 not done in one of the trainings (lack of time). Overall less time dedicated for role-plays	Session 1: completed fully; if no time then postponed to session 2 Session 2: Completed partially: both role plays were done, but no exchange of roles	Session 1: Completed fully Session 2: Completed fully

sessions. In Mexico, reach was lower compared to the other two countries, but was broadened by participating providers training their non-participating peers, which is reflected also in a relatively high percentage of screeners among providers not attending the training. The country-dependent dose delivered shows that the length of the training was adapted based on the countries' needs and familiarity of the target population with the topics (e.g. in Mexico, the depression part could be shortened as providers were already familiar with the topic through the World Health Organization's Mental Health Gap Action Programme trainings (World Health Organization, 2008), whereas for most participants in Peru the topic of alcohol screening was completely new (Kokole et al., 2021).

Fidelity of the training was related to dose; e.g. lack of time was mainly reflected in role-plays being shortened; and also to participant response: in Mexico, the videos were less well accepted in the initial trainings and were replaced with practicing with hypothetical cases. Despite assessing fidelity as part of the process evaluation, the perspective of the research and training teams when developing and implementing the trainings was aligned with the dynamic sustainability framework (Chambers et al., 2013), which suggests that quality improvement is more important than quality assurance, and considers that intervention can not be completely optimized prior to implementation,, but can instead be improved through ongoing development, evaluation and refinement in diverse contexts. From this perspective, decisions of local training teams to leave out the videos that were not well received do not represent failure to adhere to the manual, but continued refinement of the training to better fit the local context – "innovation" rather than "drift" in terminology of Bumbarger and Perkins (Bumbarger & Perkins, 2008). The activities remained aligned with theoretical background (Bandura, 1977), as suggested by Moore et al. (Moore et al., 2014), just shifting the focus from vicarious learning to enactive mastery.

Process Variables Relationship with Outcome Considerations

The developed training was shown to be successful in getting the providers to screen for risky alcohol use (Anderson et al., 2021), thus we also explored which process indicators differed between the screeners and non-screeners. Two of them stood out by relevance: dose received, and professional role.

Overall, screeners received more training (both in terms of length and number of sessions), which points to the clear importance of the dose received (within the country – between countries the amounts of training differed for reasons mentioned earlier). The dose received, however, includes attendance of booster sessions which took place during the implementation period. This means that it's possible that providers who already started screening after the first training were more likely to join the booster sessions, therefore entering a positive feedback loop (Petticrew et al., 2019) – but a more precise time series analysis would be necessary to examine these dynamics and see how that impacted the total number of screenings, which is beyond the scope of this paper. Given the difference between screeners and non-screeners in booster session attendance, booster sessions were important in all three countries, but the importance (inferred from the largest difference) was strongest in Peru, where providers had the least support from other sources and the least familiarity with the topic. In Colombia, there was organizational support, as enrolment in the study took place on the organizational level, and in Mexico, many providers had previous experience and support of health system policies, being

Figure 4. Comparison of screeners and non-screeners on process evaluation variables.

		Total	Colombia	Mexico	Peru
Outcome	% screeners among training attendees	49%	76%	53%	34%
	% screeners among non-attendees	18%	13%	29%	2%
	% screeners among all providers eligible for training	40%	69%	44%	25%
Context	Gender	50% of women and 46.3% of men screened	74.5% of women and 81.3% of men screened	58.9% of women and 38.6% of men screened	33.1% of women and 36.4% of men screened
	Age	Screeners' average age: 37.8 years; non-screeners: 39.6 years	Screeners' average age: 32.4 years; non-screeners: 31.4	Screeners' average age: 36.4 years; non-screeners: 34.7	Screeners' average age: 45.4 years; non-screeners: 44.0
	Profession	58.1% of doctors, 45.8% of nurses, 69.7% of psychologists and 36.1% other professions screened	93.1% of doctors, 81.3% of nurses, 100% of psychologists and 47.6% other professions screened	58.0% of doctors, 41.7% of nurses, 58.3% of psychologists and 40.6% other professions screened	19.2% of doctors, 20.0% of nurses, 75.0% of psychologists and 31.3% other professions screened
Implementation	Arm (short vs. standard training)	46.2% from Arm 2, 50.9% from Arm 3 and 51.2% from Arm 4 screened	72.4% from Arm 2, 81.0% from Arm 3 and 76.5% from Arm 4 screened	54.3% from Arm 2, 47.1% from Arm 3 and 57.1% from Arm 4 screened	29.4% from Arm 2, 38.2% from Arm 3 and 36.4% from Arm 4 screened
	Number of hours participated	Screeners on average more hours of training participation (2.82 vs 2.52)	Screeners: 2.84 hours; non-screeners: 2.28 hours	Screeners on average more hours of training participation (2.43 vs 2.03)	Screeners on average more hours of training participation (3.39 vs 2.89)
	Number of sessions participated	Screeners on average more sessions of training participation (1.58 vs 1.36)	Screeners on average more sessions of training participation (1.59 vs 1.25,)	Screeners on average more sessions of training participation (1.51 vs 1.26)	Screeners on average more sessions of training participation (1.69 vs 1.44)
	Participation in T1	49.7% of T1 attendees and 45.5% of non-attendees screened	82.1% of T1 attendees and 45.5% of non-attendees screened	55.4% of T1 attendees and 40.7% of non-attendees screened	32.1% of T1 attendees and 66.7% of non-attendees screened
	Participation in T2	52.2% of T2 attendees and 23.5% of non-attendees screened	/	57.9% of T2 attendees and 50.0% of non-attendees screened	45.2% of T2 attendees and 15.4% of non-attendees screened
	Booster session	60.7% of booster attendees and 41.5% of non-attendees screened	77.8% of booster attendees and 72.7% of non-attendees screened	60.5% of booster attendees and 49.0% of non-attendees screened	46.2% of booster attendees and 26.2% of non-attendees screened
Mechanisms of Impact - Participant response	Overall satisfaction - T1	Screeners: 4.57; Non-screeners: 4.52	Screeners: 4.53; Non-screeners: 4.14	Screeners: 4.53; Non-screeners: 4.71	Screeners higher satisfaction rating compared to non-screeners (4.67 vs. 4.45)
	Overall satisfaction - T2	Screeners: 4.50; Non-screeners: 4.59	/	Screeners: 4.20; Non-screeners: 4.56	Screeners: 4.80; Non-screeners: 4.61
	Overall satisfaction - B1	Screeners: 4.56; Non-screeners: 4.51	Screeners: 4.42; Non-screeners: 4.33	Screeners: 4.62; Non-screeners: 4.60	Screeners: 4.67; Non-screeners: 4.56
	Overall perceived utility - T1	Screeners: 4.66; Non-screeners: 4.63	Screeners: 4.61; Non-screeners: 4.57	Screeners: 4.63; Non-screeners: 4.76	Screeners: 4.76; Non-screeners: 4.57
	Overall perceived utility - T2	Screeners: 4.68; Non-screeners: 4.64	/	Screeners: 4.29; Non-screeners: 4.57	Screeners: 4.87; Non-screeners: 4.67
	Overall perceived utility – B1	Screeners: 4.65; Non-screeners: 4.54	Screeners: 4.58; Non-screeners: 4.22	Screeners: 4.54; Non-screeners: 4.80	Screeners: 4.79; Non-screeners: 4.59
	Other differences	Location, venue in T1 (screeners lower satisfaction rating), Exchange of experience with other participants during booster session (screeners higher perceived utility rating)	No other differences	Screeners lower satisfaction ratings with other participants at T1	Screeners higher rating on aspects of perceived utility: Learning about screening criteria for SCALA, learning about steps of brief intervention, learning about treatment options when referral is not possible, exchange of experience with other providers, practical solutions for problems

Legend:
Not significant
p ≤ 0.05
p ≤ 0.01
p ≤ 0.001

expected to include alcohol use in the patient clinical history. This also aligns with the feedback of the Peruvian providers on the importance of social support and appreciation of encountering other providers with similar interests in the training – booster sessions served as additional support in an unsupportive context.

If the dose received seemed to be similarly important across the three countries, this was not the case for the

professional role: the country-level dynamics were different, also due to different sample composition. In Colombia, most of the providers were doctors or nurses, with few psychologists, and all of those roles were more likely to screen compared to other professions. In Mexico, both doctors and psychologists were more likely to screen compared to nurses and other professions (although psychologists accounted only for a small proportion of the sample). In Peru, psychologists were more likely to screen than doctors, nurses or other professional roles. These differences perhaps reflect the differences in the country health systems and roles of professionals (specifically for substance use, but also for mental health more broadly) – e.g. in Peru substance use is often framed as part of mental health and alcohol screening is still considered as a domain of psychologists (Cavero et al., 2018), and in Colombia, PHC providers do not always consider being well equipped for dealing with mental health related ("emotional") topics (Shannon et al., 2021).

For two indicators, no differences were found: arm and participant response (satisfaction and perceived utility). No difference in the arm allocation between the screeners and non-screeners means that extra training session received by Arm 4 did not have an impact on the outcome. While this seems contradictory to the dose result above, a possible explanation is that we only looked at the first step in SCALA protocol use, which is alcohol screening, while the second training session in Arm 4 emphasized the depression part of the care pathway. Another possible explanation for the discrepancy between the two results (arm vs. dose) could also be the greater difficulty of Arm 4 training content, as providers had to master a more complex care pathway, and thus extra time in training did not translate into more practice in alcohol screening. The finding that longer and more complex intervention did not translate into more screening is aligned both with the theory (Rogers' innovation complexity theorized to lead to lower adoption (Rogers, 2003)), as well as to the evidence from primary care practice (Lau et al., 2016).

Furthermore, satisfaction with training and perceived utility for practice was not related to the outcome; possibly due to the ceiling effect, as all of the ratings were high, also rendering any differences on single items (mostly found in Peru) less practically relevant. However, a recent paper examining screening and brief intervention training effectiveness found that course completion satisfaction was not associated with immediate screening, but with the amount of screening in 12 months (Acquavita et al., 2021). Therefore, further analyses at the end of the project could be useful to associate those process indicators with the total amount of screening conducted.

Strengths and Limitations of the Study

The main strength of this process evaluation is that it employed a range of methods, combining both quantitative

Table 5. Key recommendations based on learning points from SCALA training process evaluation.

How to increase the rates of alcohol screening through provider training:

- At the individual level, dose of training is important – the more of the offered training the provider receives, the better. Increased length of the training is beneficial, unless it comes at the cost of increased complexity of the intervention: a combination of simple intervention with enough time to practice at the training is optimal. On the other hand, dose needs to be balanced with providers' availability; in our case, more than 2 h of training at the time would not be feasible.
- At country level, the amount of necessary training depended on the existing knowledge and familiarity with the topic of the providers, therefore the length should be adapted to the country context.
- Opportunity to practice, for example through role-playing, is considered helpful by the providers; thus, allocating sufficient time for it within the training session is important
- Booster sessions can serve an important role in encouraging a positive feedback loop in providers' behaviour, as they are more likely to be attended by the motivated providers who already started with implementation and need additional support. Booster sessions might be especially important in the context where less organizational or structural support is available.

and qualitative insights, which enabled a better understanding of training implementation and outcomes through data integration and corroboration through data triangulation. While we managed to collect a large amount of data in a hard-to-reach setting, there were resource and feasibility restrictions which led to some data collection limitations. For example, the only feedback received from providers was through the post-training questionnaire, which might miss some nuances of their experience. Furthermore, it was not feasible to reach all the non-attending providers, therefore data on reasons for their non-attendance had to be collected through observations from the trainers and training organizers. Finally, the number of interviews was also small and unlike to reach saturation. While not necessarily a limitation of the study per se, the local research and training teams also raised the issue of the interaction between process evaluation and training implementation – in the already scarce time available to deliver the trainings, several evaluation activities also had to be integrated, such as checking the attendance, and completing the post-training questionnaires. Another issue to be noted is the use of mechanisms of impact as a studied category – based on the MRC framework (Moore et al., 2015), we included participant response and unintended consequences as subcategories in this paper, but not variables possibly mediating the outcome (such as knowledge or attitudes), as those will be examined separately. Last, in terms of outcome, we did not make adjustments to the

amount of consultations per provider. The reason for this was that we currently only looked at whether the providers did any screening, rather than how much. Further data analysis is needed to unpack the dynamics of the amount of screening throughout the implementation of the whole SCALA project, which was beyond the scope of this paper.

Implications for Practice

Based on the results of our process evaluation, we collected a number of key learning points which might be relevant for further practice for training implementation in middle-income contexts (Table 5).

Despite the success of the training, half of the participating providers did not screen in practice, indicating that training alone was insufficient for behavior change,and other barriers apart from the lack of skills were likely impacting their screening. Training can thus be seen as a first and important step, but combination with multiple implementation strategies (such as supervision or community support) tends to produce better outcomes both on the provider and the patient level (Keurhorst et al., 2015; Rowe et al., 2018).

Conclusion

The SCALA training programme was well received by the participants and led to more providers conducting alcohol screening in primary health care in Colombia, Mexico and Peru. The training was suitable for different professional roles, but the existing health system structures meant psychologists and doctors were more likely to use the protocol after attending the training, with exact dynamics differing by country. The amount of the training received by the provider was important on the individual level, and the booster sessions were especially important in a context with less institutional support. Overall, our study showed the importance of tailoring the initial training (e.g. adapting sessions based on providers' existing knowledge) as well as ongoing refinement to better fit the local context.

Authors' contributions

EJL, HDV, PA, JR, AOD, CSS, AG, HLP, SM, FB, JMT, GNR, MP and IB contributed to the Grant Application for the study, and provided feedback on evaluation design and instrument development. AG, HLP, SM and FB developed the training package and led the country adaptation process. JMT, KM, NR, GNR, PSMA, NBA, MP and IB implemented the study at country level, collected the data and contributed to interpretation of results. DK, EJL, LM and HDV conceptualized and drafted the first version of the paper. DK analysed the data and revised the paper based on co-authors' feedback and comments. AS assisted with qualitative data analysis and interpretation. All co-authors commented on subsequent drafts of the manuscript, read and approved the final version.

ORCID iDs

Daša Kokole https://orcid.org/0000-0001-7260-8064
Natalia Bautista Aguilar https://orcid.org/0000-0003-3289-4259

Supplemental material

Supplemental material for this article is available online.

References

Acquavita, S. P., Richardson, G. B., Smith, R., Van Loon, R. A., Brehm, B., Kim, K., & Diers, T. (2021). Outcomes of an interprofessional SBIRT training program: Knowledge attainment and perceived competence for practice. *Substance Abuse*, *42*(4), 935–943. https://doi.org/10.1080/08897077.2021.1900982

Anderson, P., Bendtsen, P., Spak, F., Reynolds, J., Drummond, C., Segura, L., Keurhorst, M. N., Palacio-Vieira, J., Wojnar, M., Parkinson, K., Colom, J., Kłoda, K., Deluca, P., Baena, B., Newbury-Birch, D., Wallace, P., Heinen, M., Wolstenholme, A., & van Steenkiste, B., … Gual, T. (2016). Improving the delivery of brief interventions for heavy drinking in primary health care: Outcome results of the optimizing delivery of health care intervention (ODHIN) five-country cluster randomized factorial trial. *Addiction*, *111*(11), 1935–1945. https://doi.org/10.1111/add.13476

Anderson, P., Manthey, J., Jané-Llopis, E., Natera Rey, G., Bustamante, I. V., Piazza, M., Aguilar, P. S. M., Mejía-Trujillo, J., Pérez-Gómez, A., Rowlands, G., Lopez-Pelayo, H., Mercken, L., Kokole, D., O'Donnell, A., Solovei, A., Kaner, E., Schulte, B., de Vries, H., & Schmidt, C., … Rehm, J. (2021). Impact of training and municipal support on primary health care–based measurement of alcohol consumption in three Latin American countries: 5-month outcome results of the quasi-experimental randomized SCALA trial. *Journal of General Internal Medicine*, *36*, 2663–2671. https://doi.org/10.1007/s11606-020-06503-9

Anderson, P., Wojnar, M., Jakubczyk, A., Gual, A., Segura, L., Sovinova, H., Csemy, L., Kaner, E., Newbury-Birch, D., Fornasin, A., Struzzo, P., Ronda, G., Van Steenkiste, B., Keurhorst, M., Laurant, M., Ribeiro, C., Do Rosário, F., Alves, I., & Scafato, E., … Kolsek, M. (2014). Managing alcohol problems in general practice in Europe: Results from the European ODHIN survey of general practitioners. *Alcohol and Alcoholism*, *49*(5), 531–539. https://doi.org/10.1093/alcalc/agu043

Bandura, A. (1977). Self-efficacy; toward a unifying theory of behavioural change. *Psychological Review*, *84*(2), 191–215.

Bellos, S., Skapinakis, P., Rai, D., Zitko, P., Araya, R., Lewis, G., Lionis, C., & Mavreas, V. (2016). Longitudinal association between different levels of alcohol consumption and a new onset of depression and generalized anxiety disorder: Results from an international study in primary care. *Psychiatry Research*, *243*, 30–34.

Boden, J. M., & Fergusson, D. M. (2011). Alcohol and depression. *Addiction*, *106*(5), 906–914.

Bumbarger, B., & Perkins, D. (2008). After randomised trials: Issues related to dissemination of evidence-based interventions. *Journal of Children's Services*, *3*(2), 55–64. https://doi.org/10.1108/17466660200800012

Cavero, V., Diez-Canseco, F., Toyama, M., Flórez Salcedo, G., Ipince, A., Araya, R., & Miranda, J. J. (2018). Provision of mental health care within primary care in Peru: A qualitative study exploring the perspectives of psychologists, primary health care providers, and patients. *Wellcome Open Research*, *3*(9), 1–18. https://doi.org/10.12688/wellcomeopenres.13746.1

Chambers, D. A., Glasgow, R. E., & Stange, K. C. (2013). The dynamic sustainability framework: Addressing the paradox of sustainment amid ongoing change. *Implementation Science*, *8*(1), 1–11. https://doi.org/10.1186/1748-5908-8-117

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd Edition). Sage Publications.

Furtado, E. F., Corradi-Webster, C. M., & Laprega, M. R. (2008). Implementing brief interventions for alcohol problems in the public health system in the region of Ribeirão Preto, Brazil: Evaluation of the PAI-PAD training model. *Nordic Studies on Alcohol and Drugs*, *25*(6), 6–6. https://doi.org/10.1177/145507250802500606

GBD 2019 Risk Factors Collaborators (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *The Lancet*, *396*(10258), 1223–1249. https://doi.org/10.1016/S0140-6736(20)30752-2

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274. https://doi.org/10.3102/01623737011003255

Heather, N. (Ed). (2006). *WHO Collaborative project on identification and management of alcohol-related problems in primary health care. Report on Phase IV - Development of Country-Wide Strategies for Implementing Early identification and Brief Intervention in Primary Health Care.*

IBM Corp (2017). *IBM SPSS statistics for windows, version 25.0.* IBM Corp.

Jane-LLopis, E., Anderson, P., Piazza, M., O'Donnell, A., Gual, A., Schulte, B., Gomez, A. P., Vries, H. d., Rey, G. N., Kokole, D., Bustamente, I., Braddick, F., Trujillo, J. M., Solovei, A., Leon, A. P. d., Kaner, E., Matrai, S., Manthey, J., & Mercken, L., … Rehm, J. (2020). Implementing primary health care-based measurement, advice and treatment for heavy drinking and comorbid depression at the municipal level in three Latin American countries: Final protocol for a quasi-experimental study (SCALA study). *BMJ Open*, *10*(e038226), 1–14. https://doi.org/10.21203/rs.2.22734/v1

Kaner, E. F. S., Beyer, F. R., Muirhead, C., Campbell, F., Pienaar, E. D., Bertholet, N., Daeppen, J. B., Saunders, J. B., & Burnand, B. (2018). Effectiveness of brief alcohol interventions in primary care populations. *Cochrane Database of Systematic Reviews*, *2018*(2), 1–200. https://doi.org/10.1002/14651858.CD004148.pub4

Keurhorst, M., Heinen, M., Colom, J., Linderoth, C., Müssener, U., Okulicz-Kozaryn, K., Palacio-Vieira, J., Segura, L., Silfversparre, F., Słodownik, L., Sorribes, E., Laurant, M., & Wensing, M. (2016). Strategies in primary healthcare to implement early identification of risky alcohol consumption: Why do they work or not? A qualitative evaluation of the ODHIN study. *BMC Family Practice*, *17*(1), 1–16. https://doi.org/10.1186/s12875-016-0461-8

Keurhorst, M., van de Glind, I., Bitarello do Amaral-Sabadini, M., Anderson, P., Kaner, E., Newbury-Birch, D., Braspenning, J., Wensing, M., Heinen, M., & Laurant, M. (2015). Implementation strategies to enhance management of heavy alcohol consumption in primary health care: A meta-analysis. *Addiction*, *110*(12), 1877–1900. https://doi.org/10.1111/ADD.13088

Kokole, D., Jané-Llopis, E., Mercken, L., Piazza, M., Bustamante, I., Natera Rey, G., Medina, P., Pérez-Gómez, A., Mejía-Trujillo, J., O'Donnell, A., Kaner, E., Gual, A., Schmidt, C. S., Schulte, B., Candel, M. J. J. M., de Vries, H., & Anderson, P. (2021). Factors associated with primary health care providers' alcohol screening behavior in Colombia, Mexico and Peru. *Substance Abuse*, *42*(4), 1007–1015. https://doi.org/10.1080/08897077.2021.1903658

Lau, R., Stevenson, F., Ong, B. N., Dziedzic, K., Treweek, S., Eldridge, S., Everitt, H., Kennedy, A., Qureshi, N., Rogers, A., Peacock, R., & Murray, E. (2016). Achieving change in primary care-causes of the evidence to practice gap: Systematic reviews of reviews. *Implementation Science*, *11*(40), 1–39. https://doi.org/10.1186/s13012-016-0396-4

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'cathain, A., Tinati, T., Wight, D., & Baird, J. (2014). *Process evaluation of complex interventions UK Medical Research Council (MRC) guidance.* https://mrc.ukri.org/documents/pdf/mrc-phsrn-process-evaluation-guidance-final/.

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2015). Process evaluation of complex interventions: medical research council guidance. *BMJ (Online)*, *350*(h1258), 1–7. https://doi.org/10.1136/bmj.h1258

O'Donnell, A., Anderson, P., Newbury-Birch, D., Schulte, B., Schmidt, C., Reimer, J., & Kaner, E. (2014). The impact of brief alcohol interventions in primary healthcare: A systematic review of reviews. *Alcohol and Alcoholism*, *49*(1), 66–78. https://doi.org/10.1093/alcalc/agt170

O'Donnell, A., Schulte, B., Manthey, J., Schmidt, C. S., Piazza, M., Chavez, I. B., Natera, G., Aguilar, N. B., Hernández, G. Y. S., Mejía-Trujillo, J., Pérez-Gómez, A., Gual, A., de Vries, H., Solovei, A., Kokole, D., Kaner, E., Kilian, C.,

Rehm, J., Anderson, P., & Jané-Llopis, E. (2021). Primary care-based screening and management of depression amongst heavy drinking patients: interim secondary outcomes of a three-country quasi-experimental study in Latin America. *PloS One*, *16*(8), 1–19. https://doi.org/10.1371/JOURNAL.PONE.0255594

O'Donnell, A., Wallace, P., & Kaner, E. (2014). From efficacy to effectiveness and beyond: What next for brief interventions in primary care? *Frontiers in Psychiatry*, *5*(113), 1–8. https://doi.org/10.3389/fpsyt.2014.00113

Petticrew, M., Knai, C., Thomas, J., Rehfuess, E. A., Noyes, J., Gerhardus, A., Grimshaw, J. M., Rutter, H., & McGill, E. (2019). Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Global Health*, *4*(Suppl 1), e000899. https://doi.org/10.1136/bmjgh-2018-000899

Pinnock, H., Barwick, M., Carpenter, C. R., Eldridge, S., Grandes, G., Griffiths, C. J., Rycroft-Malone, J., Meissner, P., Murray, E., Patel, A., Sheikh, A., & Taylor, S. J. C. (2017). Standards for reporting implementation studies (StaRI) statement. *BMJ (Online)*, *356*(March), 1–9. https://doi.org/10.1136/bmj.i6795

Rogers, E. (2003). *Diffusion of innovations* (5th Edition). Free Press.

Ronzani, T. M., Fuentes-Mejía, C., Mota, D. C. B., Gomide, H. P., Ferreira, M. L., & Cruvinel, E. (2019). Brief interventions for substance abuse in Latin America: A systematic review. *Psicologia Em Estudo*, *24*(e44393), 1–16. https://doi.org/10.4025/PSICOLESTUD.V24I0.44393

Rosário, F., Santos, M. I., Angus, K., Pas, L., Ribeiro, C., & Fitzgerald, N. (2021). Factors influencing the implementation of screening and brief interventions for alcohol use in primary care practices: A systematic review using the COM-B system and theoretical domains framework. *Implementation Science*, *16*(1), 1–25. https://doi.org/10.1186/s13012-020-01073-0

Rowe, A. K., Rowe, S. Y., Peters, D. H., Holloway, K. A., Chalker, J., & Ross-Degnan, D. (2018). Effectiveness of strategies to improve health-care provider practices in low-income and middle-income countries: A systematic review. *The Lancet Global Health*, *6*(11), e1163–e1175. https://doi.org/10.1016/S2214-109X(18)30398-X

SCALA (2021). *SCALA project outputs*. https://www.scalaproject.eu/index.php/project-outputs.

Scientific Software Development GmbH (2019). *Atlas.ti 8.4 for Windows*.

Shannon, C. L., Bartels, S. M., Cepeda, M., Castro, S., Cubillos, L., Suárez-Obando, F., Williams, M. J., Gómez Restrepo, C., Uribe, M., Marsch, L., & Torrey, W. C. (2021). Perspectives on the implementation of screening and treatment for depression and alcohol use disorder in primary care in Colombia. *Community Mental Health Journal*, *57*(8), 1579–1587. https://doi.org/10.1007/s10597-021-00781-1

Smith, S. L., Franke, M. F., Rusangwa, C., Mukasakindi, H., Nyirandagijimana, B., Bienvenu, R., Uwimana, E., Uwamaliya, C., Ndikubwimana, J. S., Dorcas, S., Mpunga, T., Misago, C. N., Iyamuremye, J. D., Dusabeyezu, J. d., Mohand, A. A., Atwood, S., Osrow, R. A., Aldis, R., & Daimyo, S., … Raviola, G. J. (2020). Outcomes of a primary care mental health implementation program in rural Rwanda: A quasi-experimental implementation-effectiveness study. *PLOS ONE*, *15*(2), e0228854. https://doi.org/10.1371/journal.pone.0228854

Spagnolo, J., Champagne, F., Leduc, N., Rivard, M., Melki, W., Piat, M., Laporta, M., Guesmi, I., Bram, N., & Charfi, F. (2020). Building capacity in mental health care in low- and middle-income countries by training primary care physicians using the mhGAP: A randomized controlled trial. *Health Policy and Planning*, *35*(2), 186–198. https://doi.org/10.1093/heapol/czz138

Stokholm Bækgaard, R., Gjærevold Damhaugh, E., Mrema, D., Rasch, V., Khan, K., & Linde, D. S. (2021). Training of healthcare providers and use of long-acting reversible contraception in low- and middle-income countries: A systematic review. *Acta Obstetricia et Gynecologica Scandinavica*, *100*(4), 619–628. https://doi.org/10.1111/aogs.14127

Stoltenberg, M., Spence, D., Daubman, B., Greaves, N., Edwards, R., Bromfield, B., Perez-Cruz, P. E., Krakauer, E. L., Argentieri, M. A., & Shields, A. E. (2020). The central role of provider training in implementing resource-stratified guidelines for palliative care in low-income and middle-income countries: lessons from the Jamaica cancer care and research institute in the Caribbean and Universidad Católica in La. *Cancer*, *126*(S10), 2448–2457. https://doi.org/10.1002/cncr.32857

Suriyawongpaisal, P., Srithamrongsawadi, S., Sornsrivichai, V., & Thongtan, T. (2020). *Evaluation of Health Professional Training on a Large- Scale Implementation of Chronic Care Model in Primary Care Settings of a Middle-Income Country*. *14*(August), *14*(2), 332–342.

World Health Organization (2008). *Scaling up care for mental, neurological, and substance use disorders*. https://www.who.int/publications/i/item/9789241596206.