BMJ Open  Process evaluation protocol of a cluster randomised trial for a scalable solution for delivery of Diabetes Self-Management Education in Thailand (DSME-T)

Iliatha Papachristou Nadal,1 Chanchanok Aramrat,2 Wichuda Jiraporncharoen,2 Kanokporn Pinyopornpanish,2 Nutchar Wiwatknupakarn,2 Orawan Quansri,3 Kittipan Rerkasem,4,5 Supattra Srivanichakorn,6 Win Techakehakij,7 Nutchanath Wichit,8 Chanapat Pateekhum,2 Nick Birk,1 Elisha Ngetich,9 Kamlesh Khunti,10 Kara Hanson,11 Sanjay Kinra,1 Chaisiri Angkurawaranon 2,12

ABSTRACT

Introduction  Type 2 diabetes mellitus is a major global challenge, including for Thai policy-makers, as an estimated 4 million people in Thailand (population 68 million) have this condition. Premature death and disability due to diabetes are primarily due to complications which can be prevented by good risk factor control. Diabetes Self-Management Education (DSME) programmes provide patients with diabetes with the necessary knowledge and skills to effectively manage their disease. Currently, a trial is being conducted in Thailand to evaluate the effectiveness, defined as HbA1c<7 at 12 months after enrolment, of a culturally tailored DSME in Thailand. A process evaluation can provide further interpretation of the results from complex interventions as well as insight into the success of applying the programme into a broader context.

Methods and analysis  The aim of the process evaluation is to understand how and why the intervention was effective or ineffective and to identify contextually relevant strategies for future successful implementation. For the process evaluation, the design will be a mixed-method study collecting data from nurse providers, and village health volunteers (community health workers) as well as patients. This will be conducted using observations, interviews and focus groups from the three purposively selected groups at the beginning and end of trial. Quantitative data will be collected through surveys conducted at the beginning, during 6-month follow-up, and at the end of trial. The mixed-methods analysis will be triangulated to assess differences and similarities across the various data sources. The overall effectiveness of the intervention will be examined using multilevel analysis of repeated measures.

Ethics and dissemination  Study approved by the Chiang Mai University Research Ethics Committee (328/2018) and the London School of Hygiene & Tropical Medicine (16113/RR/12850). Results will be published in open access, peer-reviewed scientific journals.

Trial registration number  NCT03938233.

Strengths and limitations of this study

- This is the first process evaluation study of a Diabetes Self-Management Education programme for type 2 diabetes in Thailand.
- Using a mixed-method approach to triangulate from multiple data sources increases the quality of the findings.
- Quantitative data in this study relies on data that is collected from the main trial, and this may add to the trustworthiness of the process evaluation results.
- There is a risk of bias with participants who engage with the qualitative evaluation, as they may be more engaged with the intervention, and therefore, may not address problems and barriers.
- Purposive sampling of those who did not achieve diabetic control after completing the intervention and those who dropped out or declined will overcome limitations of participation bias within the qualitative evaluation.

INTRODUCTION

Type 2 diabetes mellitus is a major challenge globally, including for Thai policy-makers, as an estimated 4 million people in Thailand (population 68 million) have this condition.1 2 Premature death and disability due to diabetes are primarily due to complications (eg, ischaemic heart disease, nephropathy, retinopathy, neuropathy and foot disease) which account for twofold higher healthcare expenditure and loss of economic productivity.3 4 These complications can be prevented or delayed by good risk factor control and lifestyle interventions, though this requires a considerable degree of self-management by the patients, including adherence to multiple...
behaviours (lifestyles, medication and monitoring)\textsuperscript{5} and coping with its psychosocial impacts.\textsuperscript{6} Diabetes Self-Management Education (DSME) programmes provide patients with diabetes with the necessary knowledge and skills to effectively manage their disease.\textsuperscript{7} Furthermore, DSME programmes have been shown to be effective in over 100 studies globally and are recommended by major guidelines; however, these programmes are not widely available in Thailand.\textsuperscript{8}

**Thailand DSME trial**

The trial that is the subject of this process evaluation is the DSME in Thailand (DSME-T) trial and commenced January 2020 with completion date due April 2022. The objective of the cluster randomised control trial was to evaluate the effectiveness of the DSME-T programme under two different models of delivery: (1) a nurse-led DSME programme and (2) a community health workers-assisted DSME programme.\textsuperscript{9} The study intervention was developed from scoping for existing diabetes education programmes and culturally tailoring the programmes for a Thai audience. Stakeholder involvement (patients, caregivers, community health workers (known as village health volunteers in Thailand), clinicians and policy-makers) ensured the programme was appropriate and feasible for trial implementation. The eligibility criteria were being over 18 years of age with a new referral for type 2 diabetes management or with uncontrolled diabetes within the first 3 years of diagnosis, willing to attend education group meetings, and available for 6 months and 12 months follow-up visits. Training manuals, patient booklets and seven short films were developed to help standardise the process and contents being delivered. All materials were in Thai and were also translated into English.

**Process evaluation**

Understanding the experiences of participants is important, as it provides insight into the process of complex interventions.\textsuperscript{10} Therefore, in accordance with the Medical Research Council’s (MRC) framework (2015), a process evaluation using mixed methods will be conducted alongside the trial.\textsuperscript{10} This will take place at 6 months and towards the end of the study (at 12 months). Process evaluation is a range of contextual constructs and methodologies used to describe the multidimensional and multifactorial mechanisms underlying the effectiveness of a complex intervention.\textsuperscript{11–13} It can provide explanations that strengthen interpretation of the effect size, further understanding of best practice in clinical care and generate new questions. Therefore, using a mixed-method approach, we will evaluate what it means to first-hand experience and implement the DSME programme. In addition, views on the cultural transferability of DSME and scalability to the Thai context will be captured through interviews with the stakeholders. Multiple perspectives will be taken from nurses, community health workers and patients with type 2 diabetes mellitus.

As the overall project is a cluster trial, processes will be measured at an individual level as well as at a cluster level.\textsuperscript{11} Therefore, the control group may also be assessed to understand the “usual care” comparator.

**Aims and objectives**

The overall aim of the process evaluation is to better understand how and why the intervention was effective or ineffective and to identify contextually relevant strategies for successful implementation as well as practical difficulties in adoption, delivery, and maintenance to inform wider implementation.

Objectives are:

1. To explore how nurses and community health workers delivered all components of the DSME programme to participants as intended by the research team. This includes whether the programme was delivered as taught, how and why it varied, and to what extent changes were made in the delivery to suit the patient group (Fidelity).
2. To quantify the number of nurses and community health workers trained, as well as number of patients that participate in the programmes, including dropouts (Dose).
3. To explore the extent of patients’ participation in the DSME programme as it was intended including how they joined and completed the programme (Reach).
4. To explore how patients responded to the DMSE intervention and to what extent they experienced the programme to help better manage their diabetes (Response).
5. To identify different behavioural change techniques (BCTs) being used and potential mechanisms of action used within the DSME programme. (Clarify causal mechanisms) BCT will be classified according to the BCT V.1 taxonomy classification.\textsuperscript{14–15}
6. To explore barriers and facilitators of the DSME programme from provider, patient and policy stakeholders’ perspectives (Contextual factors) that may influence with the delivery of intervention and outcomes.

**METHODS AND ANALYSIS**

**Overall design**

The overall design will be a mixed-method study collecting data from providers, nurses and community health workers (CHWs) as well as those with type 2 diabetes (including those who dropped out and those who refused to join the programme).

**Participants and setting**

The process evaluation will include several stakeholders, that of patients, nurses and community health workers. From the main trial, patients recruited to the intervention will be from 7 primary care sites in Chiang Mai and 14 primary care sites in Lampang provinces in northern Thailand. These sites aim to manage patients with diabetes, consist of full-time nurses, with a doctor visiting weekly, and up to 15 community health workers.
All patients from the randomised sites will receive the intervention, with nurses and community health workers delivering the intervention. Overall, the total sample size of patients will be 693 (Table 1). For several components of the qualitative part of the process evaluation, out of the 21 primary care sites that are involved with the main trial, 15 of these sites will be selected through randomisation, 5 from the nurse led arm, 5 from the community health worker-assisted arm and five from the usual care arm (Table 2). In addition, trial dropped out rate will be calculated. The diabetes patients who dropped out or who declined joining the programme will be invited to take part in the process evaluation.

**Data collection**

**Quantitative data collection**

Quantitative data will be acquired through several data collection sources:

**Sociodemographic data**

Age, gender, diagnosis plus comorbidities, living situation and employment will be collected as part of the overall trial. These data will be used to examine the characteristics of participants who engage with and continue participation in the intervention, and those who do not (ie, dropouts and non-attendees). From this, programme reach will also be measured.

**Participation registration form**

The participation registration form will assess the number of sessions attended by each patient as well as number of sessions the nurses and community health workers deliver. At every DSME session, this form will be completed before commencing the session. Here, measurement of fidelity as well as attendance in the intervention will be captured. Those patients who drop out from the study, as well as those who do not attend, along with the reasons for dropout/non-attendance will be recorded on the form. This will provide input on the acceptability of the DMSE programme. The nurses and community health workers leading the DSME session will be responsible for completing the form at each contact point. For those who drop out from the study, nurses and community health workers will follow up patients with a telephone call.

**Participant’s questionnaire results**

At month 0, 6 and 12, participants will be asked to complete several standardised questionnaires to capture multiple aspects of participant’s perception and context relating to the disease. The included questionnaires are the Thai Version of Perceived Stress Scale (PSS-10),16 International Physical Activity Questionnaire (IPAQ),17 Hospital Anxiety and Depression Scale (Thai HADS),18 The Brief Illness Perception Questionnaire (B-IPQ),19 the Diabetes Management Self-Efficacy Scale (DMSES),20 The Summary Diabetes self-management Activities (SDSCA),21 the quality of life assessment (WHOQOL-Brief-Thai),22 the European Quality of life questionnaire (EQ5D),23 Chronic illness resources survey (CIRS),24 and Medical Interview Satisfaction scale (MISS-21).25

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Summary of DSME delivery in the three trial arms9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>Routine care</td>
</tr>
<tr>
<td>0</td>
<td>Individual session</td>
</tr>
<tr>
<td>6</td>
<td>Individual session</td>
</tr>
<tr>
<td>12</td>
<td>Outcome assessment</td>
</tr>
</tbody>
</table>

*Participants in the community health worker-assisted arm will additionally receive monthly contact with the community health worker either via a home visit or telephone call.

DSME, Diabetes Self-Management Education.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Semistructured interview and focus group schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview and focus group schedule</td>
<td></td>
</tr>
<tr>
<td>Time Period</td>
<td>Intervention groups</td>
</tr>
<tr>
<td>Nurse led</td>
<td></td>
</tr>
<tr>
<td>Enrolment to 6 months</td>
<td>5–8 nurses</td>
</tr>
<tr>
<td></td>
<td>5 patients focus groups</td>
</tr>
<tr>
<td>6–12 months</td>
<td>5–8 nurses</td>
</tr>
<tr>
<td></td>
<td>10 patients</td>
</tr>
<tr>
<td></td>
<td>5 patients focus groups</td>
</tr>
</tbody>
</table>
Community health workers implementation form

This form will be completed by the community health workers who provide additional monthly contact time with the participant. They will complete the form after every encounter (outside of the DSME sessions). This may be in person or via telephone. Information about the patient's health condition will be collected as well as support and advice given, topics covered and any health behavioural changes that occur. As well as dose delivered, this will provide information on fidelity. Furthermore, the community health workers will report how much contact time is needed to perform these activities. Preparation time for these activities, as well as time needed to contact other healthcare professionals and any other administration duties will also be reported.

Data storage and security

Any participants’ identifiable data collected by the Study Coordination Centre will be stored securely and their confidentiality protected in accordance with the Data Protection Act 1998.

To protect against the possibility that personally identifiable information will be accessed and used by unauthorised individuals, several security measures will be applied. Data collection devices (laptops/tablets) will be password protected. Access to electronic data on servers located at Chiang Mai University (CMU) and the London School of Hygiene and Tropical Medicine (LSHTM) will be protected using access controls including password protection; access will only be available to research personnel through the authorisation of the principal investigators. An audit trail will record activity on the main Access databases. All staff will be trained in handling of personally identifiable data. Qualitative and quantitative data will be anonymised at the earliest opportunity. Qualitative data will be used to inform intervention development and the process evaluation only; generic identifiers (eg, participant 1) will be used from the transcription stage onwards. The key linking participants' names with study IDS will be stored separately from other data in a double-locked file at the secure project office, with access restricted to appropriate study personnel. Paper consent forms will be stored similarly. Study reports, such as aggregated data in progress reports, will not contain identifying information. Project office computers will be safeguarded from theft and damage (eg, using locks, encryption and antivirus software). Fully anonymised data may be transferred for analysis to co-investigators at LSHTM and other academic and commercial partners. A secure encrypted data transfer service will be used. On request, participant records will be made available to the study sponsor.

Qualitative data collection

Qualitative data will be collected by two researchers that are not directly involved with the main trial (WJ and CAr). This is to ensure trustworthiness and external validity of the data.10 12 The Consolidated Criteria for Reporting Qualitative research checklist26 will be used throughout the qualitative methodology process, including the analysis and reporting of findings.

Data will be collected using a range of qualitative methods including one-to-one interviews with providers and focus groups with patients. Non-participatory observations and video recordings of the DSME sessions will be conducted, including video recordings of intervention delivery. In addition, patients who dropped out of the programme or who refused to participate in the programme will be interviewed on a one-to-one basis.

The following qualitative methods will be used for data collection:

Interviews

Semistructured interviews consisting of open-ended questions will be conducted with nurses, community health workers and patients in all groups (the nurse-led group, community health worker-assisted group and control group). The total number of providers that will be interviewed are described in table 2. All nurses and community health workers involved in delivering the DSME session will be invited via email or telephone to take part in the interview. Patients will be invited to take part in the interview at their follow-up visit at the end of the trial. Furthermore, a small subset of patients who dropped out of the programme or those who refused to participate in the programme will be interviewed on a one-to-one basis. The aim is to interview at least one patient who was able to control their diabetes and one patient who could not from each selected site. Interviews will be arranged at a time and place of convenience to the participant (eg, the place of their usual clinic). A topic guide will be used to explore the experiences of implementing the intervention, barriers and facilitators of delivering the DSME programme, and continuity of delivery. Specifically, questions will include providers’ opinions on the training received; how they prepared to deliver the programme; the materials available to them, (ie, training manuals, booklets for patients and education videos); changes they made during the implementation process. Questions will also include patients’ opinions on their experience with DSME sessions, how they applied what they learnt to improve their self-management skills and support they received. As both nurses and community health workers are spread across each different primary care site, collecting the data in the form of a one-to-one interview at the site where the provider is situated is a more convenient method. Each interview will last between 45 min and an hour.

Focus groups

Focus groups consisting of up to 6 patients per group will be conducted in the intervention group (the nurse-led group and community health worker-assisted group). It is anticipated that patients will be recruited to participate in the focus group at the point of attending a DSME session. Those who agree will remain behind after the session to take part in the focus group. Participants who
dropped out or refused to participate in the trial will be invited to a separate focus group, one for non-attendees and one for those who refused to participate. A topic guide will be used to explore the experiences of the DSME programme, including barriers and facilitators of receiving and applying the DSME programme to their everyday lives. The extent of patients’ participation to the programme will be explored as well as their views on continuity and scalability of the programme. Each focus group meeting will last up to 90 min. A 5–7 focus groups will be conducted across the 7 sites with 30–42 patients participating overall.

The interviews and focus groups will provide insight into the acceptability of the DSME programme from both providers’ and patients’ perspectives. These interviews will also provide further insight into dose delivered and received, further modifications needed on the delivery of the programme, and potential improvements for a scalable DSME programme. All participants will be invited back for a follow-up interview towards the end of the intervention (at 9–12 months).

Non-participatory observations and videorecordings
Non-participatory observations and videorecordings of DSME sessions will be conducted in the intervention groups (nurse led and the community health worker assisted). At least 6 out of the 14 interventions sites will be randomly selected to be approached and to conduct the observations, resulting in at least three sites for each intervention group. A member of the research team will spend time providing observations across both intervention arms. These observations will consist of at least 16 videorecordings and audiorecordings of the DSME sessions across at least six different sites (three in each intervention). Field notes will also be taken during the recording. The allocated researcher(s) will spend time observing providers and patients within the DSME sessions to understand the implementational processes. The field notes will help to generate rich descriptions of the delivery of the individual session. Field notes will be kept and coded for emerging themes.

The videorecording will allow detailed insight into patient–provider interaction and engagement that otherwise may be missed. It will also allow researchers to identify different BCTs being used to deliver the DSME. For this research, we will use the BCT V.1 taxonomy classification of 16 commonly used BCTs (clusters), further defined into 93 BCT labels. The DSME training programme has been designed using this BCT taxonomy, and therefore, there will be a strong distinction across each BCT when mapping these to each video recording. Each video recording will be up to 2–3 hours long, and therefore, a rich amount of data will be collected. By identifying these techniques, the programme can be linked to different intervention functions which can ultimately help to provide insight on what necessary intervention techniques are required to bring effective change among the participants. In addition, data will be collected on the number of sessions participants attended or reasons for non-attendance at each session.

Document review
Researchers will also review existing locally relevant documents related to the number of new cases of diabetes being registered within the country from the Thailand Health Data Centre website and Thailand health survey reports. We will look at how resources are being allocated for DSME programmes from documents that record the content of meetings between researchers and stakeholders at the planning process of the DSME-T trial as well as other relevant documents that will provide insights on the potential issues related to implementation and scalability of the DSME-T.

An overall summary of the different methods of data collection and key objectives is given as table 3.

Data analysis
Quantitative data analysis
As stated in the main trial protocol, the effectiveness of the intervention will be evaluated using multilevel analysis of repeated measures. The main outcome is changes in hemoglobin A1C (HbA1c). The analysis will account for clusters effect within the study.

The quantitative data collected will act as the process variables that will account for dose, reach and fidelity. These will be taken from the sociodemographic data, participant registration form, participant questionnaire results, community health worker implementation form and documents from DSME sessions. All data will be entered and analysed using SPSS version 22 or STATA version 15. The number of DSME sessions done by each study site, the total number of participants in each site, the number of patients in each session and the frequency of CHWs follow-up visits are numbers representing fidelity quantitatively. Fidelity will be analysed descriptively along with its qualitative analysis counterpart.

The reach of the intervention (ie, did the intended population participate in the study) will be assessed descriptively by calculating the proportion of patients included in the intervention over the entire target population within the study site. To assess for potential factors that might associate with the reach, sociodemographic variables of those who participate vs drop-outs and non-attendees will be compared using $\chi^2$ for categorical variables or Analysis of variance (ANOVA) for continuous variables. Exploring changes in participants’ questionnaire scores would provide information on how the participants responded to the intervention. These questionnaires include B-IPQ, DMSES, IPAQ, and the Summary of Diabetes Self-Care Activities Questionnaire (SDSCA). These questionnaires represent educational target area of different components of DSME as follow. SDSCA represents diet, nutrition, physical activity and exercise. IPQ represents physical activity and exercise. IPQ and DMSES represent stress management and mental health. Exploring changes in the questionnaires
would provide the effect of each component of DSME on participant’s behaviour. Potential contextual factors that may affect effectiveness will also be included in the analysis. These include CIRS,24 MISS-21,25 PSS-10,16 HADS,18 WHOQOL-BREF (a quality-of-life assessment),22 and the EQ5D.23

Qualitative analysis
The MRC guidance for process evaluation will act as a framework matrix for the qualitative data in the analysis phase. Qualitative data from the interviews, focus groups and observations will be transcribed, and transcripts will be analysed using NVivo software. A thematic analysis will be conducted to identify recurrent and unique themes across the different participation groups.29 This approach is inductive (themes emerge from the data and are not imposed on it by the researcher) and iterative (data collection and analysis occur simultaneously).30 Data will be coded and classified by at least two members of the research team according to themes that arise out of the data. Comparative analysis will also be carried out; this method allows data from different participants to be compared and contrasted. Deviant cases will be actively sought throughout the analysis, and emerging ideas and themes will be modified in response.31

Table 3 Methodological overview of process evaluation

<table>
<thead>
<tr>
<th>Summary</th>
<th>Objectives</th>
<th>Method for extracting objectives</th>
<th>Questionnaires</th>
<th>Video and audio recording the session</th>
<th>Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fidelity</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was each component of the intervention provided as intended?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>How are resources being mobilised/allocated to achieve the intervention?</td>
<td>x (nurses and community health workers)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Dose (frequency of the intervention delivered and received)</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Reach</td>
<td></td>
<td>x</td>
<td>(evaluate biases of the estimated proportion, eg, hidden population)</td>
<td>x (estimated proportion)</td>
</tr>
<tr>
<td></td>
<td>How has the intervention reached the participants?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The estimated proportion of the intended participants who participate in the intervention</td>
<td>x (estimated proportion)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Response (How participants respond to the intervention)</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>During the session</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>How what they learnt are related to DM control</td>
<td>x (B-IPQ, DMSES)</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>How did they apply what they learnt?</td>
<td>x (IPAQ, SDSCA)</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>Clarify causal mechanisms</td>
<td>x (IPAQ, B-IPQ, DMSES, SDSCA)</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Component of DSME</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>What BCTs are used by providers when delivering the intervention</td>
<td>x (IPAQ, B-IPQ, DMSES, SDSCA)</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Contextual factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Resource support from stakeholders</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Family and community support</td>
<td>x (CIRS)</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session/learning environment</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Patient perception regarding the disease/illness</td>
<td>x (B-IPQ, DMSES)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Satisfaction to health system</td>
<td>x (MISS-21)</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Baseline characteristics of the participants</td>
<td>x (PSS, HADS, WHOQOL, EQ5D)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BCT, behavioural change technique; B-IPQ, Brief Illness Perception Questionnaire; CIRS, Chronic Illness Resources Survey; DMSES, Diabetes Management Self-Efficacy Scale; DSME, Diabetes Self-Management Education; EQ5D, European Quality of Life questionnaire; HADS, Hospital Anxiety and Depression Scale; IPAQ, International physical activity questionnaire; MISS-21, Medical Interview Satisfaction Scale; PSS, Thai Version of Perceived Stress Scale; SDSCA, The Summary Diabetes Self-Care Activities; WHOQOL, The WHO Quality of Life.
Machine learning

A pooled qualitative dataset of transcripts will be extracted from video and audio recordings. From these transcripts, the data will be assigned labels according to BCT categories by the researchers to enable the use of supervised machine learning methods for text data. This labelled data will be used in training a multiclass text classification model. Using a ‘bag of words’ approach, an algorithm will be developed to classify sentences into the appropriate BCT category or to categorise ‘non-BCT’ statements which do not fit into any of the taxonomy categories (eg, statements not related to the intervention or diabetes). To this end, machine learning techniques including Naive Bayes and Support Vector Machines will be fit to this labelled subset of the data. We will use cross-validation to assess the predictive performance of the models within this training subset. Once appropriate prediction accuracy is achieved, the best-performing classification model will be deployed on the remaining interview text data to automate the classification of statements into each of the BCT categories. To this end, we will be able to reduce the time required to quantify the frequency with which techniques from each BCT category are mentioned by participants in the focus groups. Models will be developed using R or Python.

Integrating results of the analysis

The findings of the process evaluation will be completed before the main trial and will be reported independently of the main trial. The mixed-methods analysis will be triangulated using the Good Reporting of a Mixed Methods Study framework, assessing for differences and similarities across the various data sources. The overall findings will be synthesised to demonstrate what worked and what did not work across the various components of the intervention. This will inform a better understanding of the programme’s implementation beyond the duration of the trial.

Patient and public involvement

The overall DSME-T trial was developed with input from a wide range of international experts and people with lived experiences. These stakeholders consisted of healthcare professionals such as nurses, doctors and community health workers within the northern provinces of Thailand (Chiang Mai and Lampang), as well as people with diabetes and their family members. Policy-makers and academic experts not directly involved with the trial have also had input into the trial development. Dissemination of results to study participants will be delivered through the patient organisation and social media.

Author affiliations

1Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK
2Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand
3ASEAN Health Institute for Health Development, Mahidol University, Salaya, Nakhon Pathom, Thailand
4Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand
5NCD Center of Excellence, Research Institute for Health Sciences, Chiang Mai University, Chiang Mai, Thailand
6Department of Disease Control, Ministry of Public Health, Bangkok, Nonthaburi, Thailand
7Lampang Hospital, Lampang, Thailand
8Surat Thani Rajabhat University, Surat Thani, Thailand
9Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
10Diabetes Research Centre, University of Leicester, Leicester, UK
11Department of Global Health and Development, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, UK
12Global Health and Chronic Conditions Research Group, Chiang Mai University, Chiang Mai, Thailand

Contributors IPN, CA, WI, NB, EN, SK and CAn were involved in the conception and design of the process evaluation; IPN, KP, OQ, KR, SS, WT, CP, NW, KK, KH, SK and CAn were involved in the conception and design of the main trial; IPN, CA, NB and CAn drafted the manuscript; NW, EN, SK, KP, OQ, KR, SS, WT, NW, KK and KH critically revised the manuscript. All authors approve the final version and agree to be accountable for all aspects of the work.

Funding This study is supported by UK Medical Research Council (MRC) grant number (MR/R020876/1) and the Thailand Research Fund (TRF) grant number (DBG6180007). This study was also partly supported by Chiang Mai University.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Chaisiri Angkurawaranon http://orcid.org/0000-0003-4206-9164

REFERENCES

34 Rennie J, Rifkin R. Improving multiclass text classification with the support vector machine. Cambridge, MA: Massachusetts Institute of Technology (USA), 2001.