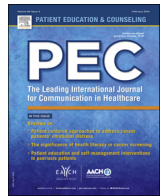




Contents lists available at ScienceDirect

Patient Education and Counseling

journal homepage: www.elsevier.com/locate/pateducou



Review

Assessment of implementation fidelity in diabetes self-management education programs: A systematic review

Louise Schinckus*, Stephan Van den Broucke, Marie Housiaux on behalf of the Diabetes Literacy Consortium¹

Psychological Sciences Research Institute, Catholic University of Louvain, Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:

Received 21 October 2013
Received in revised form 19 March 2014
Accepted 5 April 2014

Keywords:

Diabetes
Self-management education
Implementation fidelity
Literature review

ABSTRACT

Objective: As diabetes requires extensive self-care, self-management education is widely recommended to enhance the effectiveness and reduce the costs of treatment. While a variety of diabetes self-management (DSM) programs are available, the conditions for their effective implementation are not well documented. This paper reviews the literature on implementation fidelity (IF), the degree to which programs are delivered as intended, as a factor influencing the effectiveness of diabetes education.

Methods: Medical, psychological and educational research databases were searched to identify published studies on diabetes education describing the implementation process. Studies detailing the intervention adherence/fidelity/integrity were included to assess the way key elements of IF were addressed.

Results: From an initial 418 abstracts, 20 published papers were retained for an in-depth analysis focusing on the components of IF. Intervention content was mainly assessed through observation, whereas intervention dose was more often assessed through self-report measures. Only one study addressed the relationship between IF and intervention effectiveness.

Conclusion: Despite the importance of IF to achieve program outcomes, IF of DSM programs remains largely under-investigated.

Practice implications: The results of this review suggest that reports on DSM education should systematically describe how the program was implemented. The impact of IF on program outcomes needs further investigation.

© 2014 Elsevier Ireland Ltd. All rights reserved.

Contents

1. Introduction	000
1.1. Diabetes self-management programs	000
2. Implementation fidelity	000
3. Method	000
3.1. Literature search and selection	000
3.2. Analysis	000
4. Results	000
4.1. Existing reviews and theoretical papers	000
4.2. Empirical studies	000
4.2.1. Description of the interventions	000
4.2.2. Conceptualization of implementation fidelity	000
4.2.3. Assessment of implementation fidelity	000
4.2.4. Level of implementation fidelity	000
4.2.5. Implementation fidelity and program outcomes	000

* Corresponding author at: Psychological Sciences Research Institute, Catholic University of Louvain, 10 Place Cardinal Mercier, 1348 Louvain-la-Neuve, bte L3.05.01, Belgium. Tel.: +32 10 47 80 68; fax: +32 10479774.

E-mail addresses: louise.schinckus@uclouvain.be, schinckuslouise@gmail.com (L. Schinckus).

¹ See Appendix A.

5.	Discussion and conclusion	000
5.1.	Discussion	000
5.2.	Study limitations.	000
5.3.	Conclusion.	000
5.4.	Practice implications.	000
	Acknowledgements	000
	References	000

1. Introduction

Diabetes mellitus (DM) is one of the most challenging health problems of our time. As one of the most common non-communicable diseases globally, it is the fourth or fifth leading cause of death in high-income countries, and is rapidly becoming epidemic in many developing and newly industrialized countries. In 2011, the number of persons suffering from diabetes was estimated at 366 million worldwide, almost 50% of whom are unaware of their condition [1]. By 2030 this number is expected to rise to 552 million. As treatment and complications are costly, diabetes care takes up between 5 and 15% of total health expenditure, depending on the country [2].

Of the three main types of DM (type 1 diabetes caused by the body’s failure to produce insulin, afflicting mainly children and teens; type 2 diabetes resulting from insulin resistance related to aging, sedentary lifestyle, poor diet, genetic influence, and obesity; and gestational diabetes occurring in pregnant women without a previous diagnosis of diabetes), type 2 diabetes is by far the most common, making up approximately 95% of all DM cases. Its prevalence is rising rapidly and is expected to increase in the coming years as a result of aging populations, increasing urbanization, obesity, dietary changes, reduced physical activity, and other unhealthy behaviors [3]. As early diagnosis and appropriate management of type 2 diabetes significantly increases the chances of preventing harmful and costly complications, the care for patients with this type of diabetes focuses strongly on the disease management, and especially on self-management by patients.

1.1. Diabetes self-management programs

Because diabetes requires extensive self-care, the capacities of patients to manage their own illness and care process are considered as a key determinant of treatment outcomes [4]. To enhance these capacities, education of diabetes patients is widely recommended and carried out [5]. Diabetes self-management (DSM) education is defined as the process of teaching persons with diabetes to manage their illness and treatment by providing them with the knowledge and skills that are needed to perform self-care behaviors, manage crises, and make lifestyle changes [6,7].

The above definition allows for a variety of educational approaches to DSM. Educational interventions range from brief instructions by physicians, nurses, or dieticians to more formal and comprehensive programs [6]. A meta-analysis by Brown [9] showed a significant shift in the types of education programs over time. In the 1960s and 1970s, DSM interventions were brief, individually oriented, and mostly delivered in the hospital setting by a nurse or a dietician. From 1980 onwards, more specific programs have been set up for diabetes patients and their relatives, whereby health care professionals with different disciplinary backgrounds educate patients in their own domain of expertise. In addition to individual education of patients, more cost-effective alternatives such as group-based education [10], information technology (IT)-based education [11] and self-help programs or support groups have been developed [12].

2. Implementation fidelity

The success of a diabetes education program not only depends on the strategy and methodology that is used, but also on the quality with which it is implemented. Given the demonstrated efficacy of existing strategies to improve glycemic control, increase physical activity and improve diet, the main public health challenge is not to find new efficacious treatments, but to implement the proven programs with consistency and efficiency [14].

A key element of the quality of implementation is its *fidelity*, or the degree to which the intervention is delivered as intended [15]. There are several reasons why implementation fidelity (IF) merits attention [16]. (a) Without information about the program delivery, the absence of significant effects may lead to a false attribution of the lack of an intervention’s effectiveness to the shortcomings of the intervention itself, when it could have resulted from poor implementation. This phenomenon has been dubbed the “type III error” [17]. (b) Information about IF can help one understand *why* an intervention succeeded or failed. (c) Assessing IF can help to identify which components have been adapted to meet the specific needs of the health system and its patients, and how these adaptations influenced the outcomes. (d) Information on IF can help to assess the future feasibility of implementing the intervention, thus serving formative in addition to summative evaluation purposes.

There are different approaches to assess the IF [18]: (a) According to the *critical components approach*, a program is composed of several critical components, and the outcomes of the program depend on their presence or absence. To assess IF, researchers need to verify whether all the critical components have correctly been implemented. In this perspective, tools to assess IF look like a “fidelity index” that are very specific to a particular program [e.g. 19]. (b) The *structure and process approach* follows the logic of critical components and tries to characterize each component as “structure” or “process”. Structural components can be related to resources and framework for service delivery, whereas procedural components can be related to roles and behaviors, or to the way in which services are delivered. According to this approach, IF depends on both the composition (structure) and the human interaction that occurs during the delivery (process) [e.g. 20]. (c) The *dimensional approach* considers IF as a multidimensional concept whereby each dimension can be assessed separately. Although certain dimensions such as adherence, exposure and quality are commonly mentioned [e.g. 21,22], other dimensions vary between authors. These three different approaches may lead to very different IF measurements. As such, the critical components approach makes it possible to assess very specific aspects of the intervention, whereas the dimensional approach may allow researchers to compare the IF of different kinds of interventions.

Of the various theoretical models proposed to consider IF, the model developed by Carroll et al. [15] is the most comprehensive [23]. This model has a dimensional approach but integrates the notion of critical components in one of its dimensions, notably the content of the intervention. The principal concept in this model is

the *adherence*, which refers to the degree to which the active ingredients of the intervention have been delivered to the participants with the planned frequency, duration and intensity. Adherence is operationally defined by four components: (a) the *content* of the intervention (was the full content delivered to the participants?), (b) its *frequency* and (c) *duration* (was the intervention delivered with the frequency and duration prescribed by the developers?), and (d) *coverage* (have all the persons who should have participated in the intervention actually done so?). The combination of the frequency, duration and coverage of the intervention are referred to as the *intervention dose*. The level of IF may be moderated by four interrelated variables: (a) *Intervention complexity* refers to the nature and comprehensiveness of the intervention, whereby an intervention is more complex if several providers are involved, if it comprises several sessions, and if there are several groups of participants [24,25]; (b) *Facilitation strategies* such as a manual, training and feedback help to optimize and standardize the fidelity of the implementation; (c) The *quality of delivery* refers to the dedication of the individuals who are responsible for delivering the intervention. Durlak and DuPre [26] suggested that provider characteristics such as perception of the need and the benefits of the intervention, self-efficacy, and skills may affect the fidelity of a program's implementation. (d) *Participant responsiveness* refers to the fact that higher IF is achieved when the participants are more enthusiastic about the intervention. Hasson [23] suggested two additional moderators, notably (e) *Recruitment*, which involves the procedure for selecting and recruiting participants, the reasons for non-participation, and the presence or absence of specific participant subgroups; and (f) *Context*, which refers to the culture and the organizational structure in which the intervention takes place (e.g. positive working climate, norms to change, share decisions, communication).

In recent years, efforts have been undertaken to measure and optimize IF in terms of the above dimensions or similar ones for a range of educational programs. Existing reviews of studies of DSM using these measures indicate that program guidelines are often poorly implemented [7], and that studies of self-management interventions seldom provide sufficient detail of the implementation to consider replication or application to other clinical settings [9]. However, a systematic review of IF applied to self-management education programs for patients with diabetes has thus far not yet been reported. The present paper reports on the results of a systematic literature review of this subject, with a view to clarify how the IF is operationalized and assessed in existing DSM education programs and how it affects program outcomes. The results of such a review have important bearings for public health practice and policy.

3. Method

3.1. Literature search and selection

A search of the electronic databases PsycInfo, PsycArticle, MEDLINE, PubMed, and Google Scholar was undertaken in January and February 2013, using the search terms “self-management” AND “diabetes” in the title and the abstract, AND (“intervention adherence” OR “implementation fidelity” OR “intervention integrity” OR “intervention fidelity”) in full text. Preliminary searches showed that other search terms such as “intervention compliance” did not lead to identifying relevant articles. To select relevant publications, the following inclusion criteria were used: (1) The primary focus of the study had to be on self-management education for diabetes; and (2) The issue of implementation quality or fidelity had to be raised. As the number of empirical studies meeting these criteria was very low, theoretical articles and interventions for people at risk of type 2 diabetes (i.e., pre-diabetes) were also

included. To be included in the review, interventions for people at risk had to focus on at least one of the seven self-care behaviors recommended by the AADE [8] for DSM education, and raise the issue of IF. Study protocols were excluded from the review as they did not report any results. For practical reasons, only articles written in English were retained. No restriction was used concerning the publication date. The oldest article was a theoretical article by Glasgow published in 1995. All the empirical studies were published between 2003 and 2013. References from relevant studies were also examined to find additional titles.

When applying the above search procedure, and after elimination of duplicate articles, a total of 418 articles were identified. Twenty articles met the inclusion criteria and were selected for further analysis. Of these, 5 were literature reviews or theoretical articles, and 15 concerned empirical studies. The search process is summarized in the flow chart shown in Fig. 1.

3.2. Analysis

Articles that met the inclusion criteria were first separated into two groups: existing literature reviews and theoretical papers, and empirical studies. One protocol was also included in the review as it had been used to collect data and some results about IF were presented in the article concerned [27]. Empirical studies were evaluated for completeness of information by assessing whether the following information was provided: (a) description of one or more self-management intervention(s); (b) indication of type 2 diabetes patients as the target group for the intervention(s); (c) reference to the concept of IF; (d) presentation of at least one approach to assess IF; (e) presentation of results related to IF. Two reviewers assessed independently whether this information content was provided in each article. Inter-reviewer agreement for each item was assessed using Cohen's kappa, yielding a perfect agreement ($\kappa = 1$) for the identification of the target group, and a substantial agreement ($\kappa = 0.68$) for the identification of a self-management intervention. For the three items related to IF conceptualization, assessment approach and results, the agreement was fair ($\kappa = 0.40$). Consequently, for these variables a consensus procedure between both reviewers was used to determine if the information was available.

In a second step, in-depth reviews of the full text were performed on all selected articles, performing separate analyses for existing literature reviews and theoretical papers and for empirical studies. For the first, the content analysis focused on the extent to which IF was addressed. For the empirical studies, the analysis focused on five specific questions: (a) How was IF conceptualized? (b) How was IF measured? (c) How were the different components of IF assessed? (d) What was the level of IF? And (e) was there a relationship between IF and intervention outcomes? For the first question, the model of Carroll et al. [15] was used as a reference to check if the provider's adherence to the content, duration, frequency, and coverage of the intervention were included. For the second and third questions, Schoenwald et al. [22] categorization of measures of IF was applied, distinguishing between (a) observation (direct or by audio/video tape), (b) self-report measures by the provider (through questionnaires or interviews), and (c) self-report measures by participants (through questionnaires or interviews).

4. Results

4.1. Existing reviews and theoretical papers

Five literature reviews and theoretical papers focused on self-management interventions for diabetes patients and mentioned IF [28–32]. In the theoretical articles, IF is often considered in a general way. Glasgow [28] mentions IF in his conceptual model for

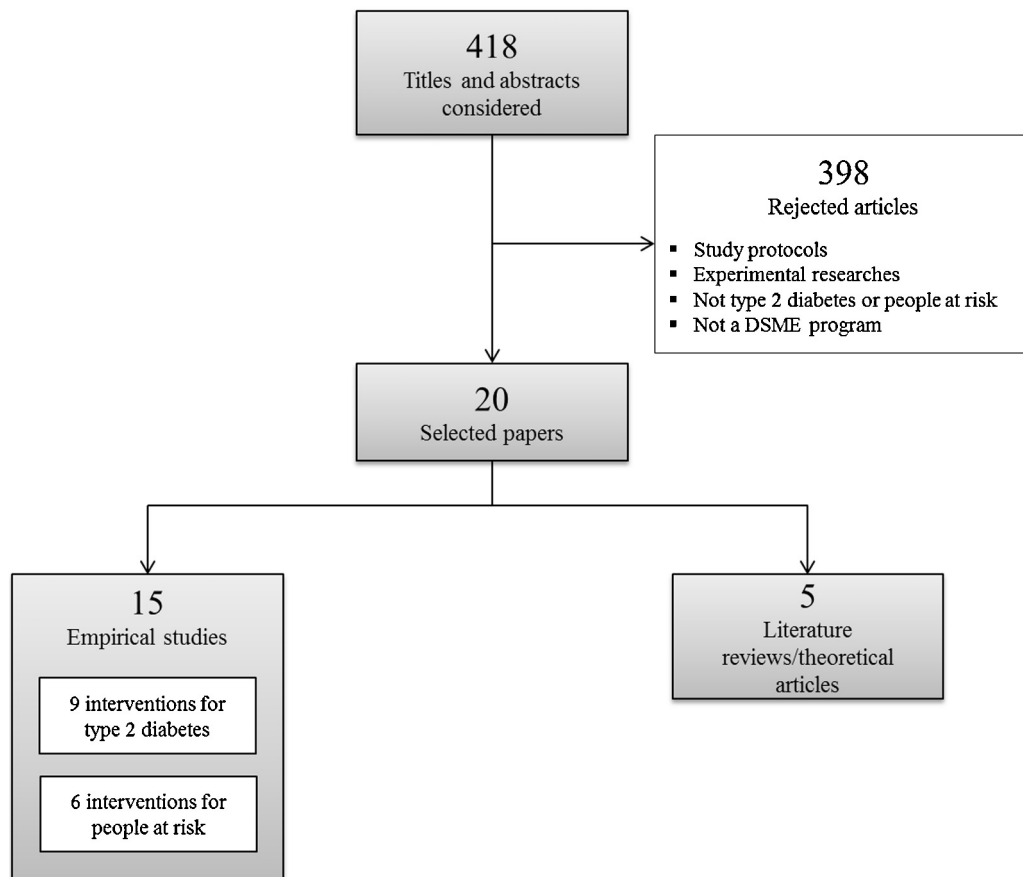


Fig. 1. Flow chart of selection process for the review.

DSM education, which in addition to the social environment, contextual factors, self-management behavior, and health and quality of life outcomes also focuses on the interaction between the patient and health care providers. The inclusion of this interaction component is expected to increase providers' adherence to recommended guidelines for diabetes care. In a similar way, Hayes et al. [31] propose coaching by nurse practitioners as a means to overcome poor adherence to guideline recommendations: providers who receive direct feedback in a coaching relationship are expected to stay motivated and committed to the coaching relationship. McAndrew et al. [30] discuss how a common sense model of self-regulation can be used to design interventions for chronic illness management, giving DSM as an example, and mention IF as a factor that needs to be considered before planning and implementing the intervention.

While these examples show that theoretical models of DSM pay attention to the providers' adherence to guideline recommendations, IF is only mentioned explicitly in two reviews of DSM interventions. Hunt [32] presents two behavioral coding systems to evaluate treatment fidelity to motivational interviewing principles applied to diabetes. The first (Motivational Interviewing Skill Code) considers patients' and providers' behaviors, and the second (Motivational Interview Treatment Integrity) focuses on the providers' performance only. Leeman, Jackson [29] look at barriers to implementation in the way intervention studies are reported in journal publications, and conclude that most published studies do not provide enough information on essential aspects such as the frequency and duration of patient contacts or the training of providers to allow potential users to adopt and implement an intervention in practice. Moreover, in only two out of the 46 studies included in their review, more intervention material was made available upon request.

If IF is only sparsely addressed in existing reviews and theoretical papers, there are even fewer attempts to operationally define the concept. Only two reviews and one theoretical article present an operationalization of IF. The scales described by Hunt [32] include *content*, as well as *participant responsiveness*, and *quality of delivery* as indicators of the IF. Leeman et al. [29] operationalize IF by considering the *frequency* and the *duration* of interventions. McAndrew et al. [30] distinguish between the assessment of IF in terms of *how treatment is delivered*, *how it is received*, and *how it is enacted*. In all these cases, the operationalization of IF is limited to a small number of dimensions, and pays no attention to moderating factors as proposed in the model of Carroll et al. [15].

4.2. Empirical studies

Fifteen empirical papers addressing IF in the context of DSM were identified in the literature. Of these, 9 concern type 2 diabetes patients as the main target group, and 6 focus on pre-diabetes patients and/or people at risk. Table 1 summarizes the information about these 15 studies, which are all included in this review. In the next sections, we will present (a) the description of each intervention, (b) the way in which IF and its potential moderators were conceptualized, (c) the way in which IF and its components were operationalized, (d) the level of IF achieved, and (e) the relationship between IF results and intervention outcomes.

4.2.1. Description of the interventions

Nine of the 15 studies [27,33–40] provide complete information about the program, the conceptualization and assessment of IF, and the results of the intervention. Two studies describe the program and offer a conceptualization of IF, but do not provide any

Table 1
Information provided in selected empirical studies.

	N (number of patients/participants)	Self-management	Diabetes type II	IF operationalization	IF measure	IF results	RCT
Di Loreto et al. [33]	340	✓	✓	✓	✓	✓	✓
Perrin et al. [34]	14 health centers	✓	✓	✓	✓	✓	Randomized control trial of health centers
Taub [35]	2287	✓	✓	✓	✓	✓	Subsample analyses
Huizinga et al. [27]	164	✓	✓	✓	✓	✓	✓
Griffin et al. [36]	8159 participants spread over 12 sites	✓	At risk	✓	✓	✓	Program assessment
Whittemore et al. [37]	58	✓	At risk	✓	✓	✓	Randomized controlled trial of 4 sites
Castro et al. [38]	181	✓	At risk	✓	✓	✓	✓
Lakerveld et al. [39]	490	✓	At risk	✓	✓	✓	✓
Rothschild et al. [40]	144	✓	✓	✓	✓	✓	✓
Siminerio et al. [41]	104	✓In Chronic care model	✓	✓	✓	✓	Non-randomized controlled trial
Haltiwanger et Galindo [42]	1 (case study)	✓	✓	✓	✓	✓	Case study
Schouten et al. [43]	1 125	✓	✓	✓	✓	✓	Non-randomized trial
Cherrington et al. [44]	16 programs	✓	✓	✓	Interviews with intervention developers	✓	Program identification and assessment
Pagoto et al. [45]	118	✓	21% type II and 79% at risk	✓	✓	✓	Non-controlled trial
Kramer et al. [46]	93	✓	At risk	✓	✓	✓	Non-randomized trial
<i>n</i>	15		9	11	12	11	5

information on the way in which the concept was assessed [41,42]. Another study, describing a quality improvement intervention to improve providers' adherence to diabetes standards [43], has assessed IF and report the results, but does not detail the conceptualization of IF on which the assessment was based. Finally, three publications describe studies involving IF of DSM interventions, but do not provide information about either the conceptualization or the assessment of the concept [44–46].

The self-management programs considered in these 15 studies all focus on a range of self-management behaviors, including glycemic control, medication adherence, maintaining a healthy diet, physical activity, smoking cessation, reducing alcohol consumption, or, more generally, increasing self-care behaviors or improving knowledge concerning diabetes and its complications. The approaches for enhancing these behaviors include: (a) *individual approaches*, such as face-to-face counseling sessions [37], therapy sessions [42], face-to-face counseling sessions followed by phone contacts [33,36,39], and phone contacts assessing the patients' adherence to treatment and providing assistance to overcome barriers to self-management [27,38]; (b) *group interventions* supervised by trained educators [34,36,41,45,46]; and (c) *peer support interventions* with a community-based approach involving persons from the target group as providers [40,44]. Two articles [35,43] do not specify whether counseling has been given individually or in groups. The providers involved in these 15 studies were nurse practitioners [37,39], nurses and dietitians [27], dietitians, exercise physiologists and psychologists [45], physicians [33], trained health educators [36], certified diabetes educators [41], clinic staff [34,43] unspecified health professionals [35,38,42,46] or members from the target population [38,40,44]. The duration of the intervention is mentioned for 11 studies [27,33,34,36,38–42,45,46] and varies between 8 weeks [42] to 24 months [27,33,40]. Sample sizes vary between $n = 1$ [42] for a case study and $n = 8159$ [36]. Five of the studies involved randomized controlled trials [27,33,38–40], and for three others [34,37,43], randomization was done at the level of health centers. Intervention effects were assessed through self-monitoring of blood glucose, HbA1c, medication adherence, diet and

physical activity, body mass index, blood pressure, and average energy expenditure.

4.2.2. Conceptualization of implementation fidelity

Among the 15 selected empirical articles, 12 provide specific information about the conceptualization of IF. A content analysis of these articles yield the information summarized in Table 2.

As this table shows, the most commonly used indicator of IF is *adherence*, most often defined operationally in terms of the *content* of the intervention [27,33–42]. This is followed by *frequency* [33,36,37,42], *duration* [36–38] and *coverage* [35–37]. In three studies, the quality of the program delivery was used as an indicator of IF [38–40], rather than as a potential moderator. Four articles do not specify how IF have been conceptualized [43–46]. Some articles introduce components that are not considered in Carroll et al.'s [15] model, such as the *number of adaptations* (amount of changes introduced to the original protocol) [36], and *program differentiation* (implementing the unique features of the intervention to distinguish it from other interventions) [42].

In terms of the potential moderators of IF, several studies refer to *facilitating strategies* to enhance the implementation quality: (a) providers' trainings were set up in the studies described in 11 articles [34,36–42,44–46]; and (b) intervention protocols and/or technical assistance were available to support providers during the intervention delivery in 9 studies [27,33,34,36–38,40,43,46]. One article reported on the positive impact of regular control with audiotaping of IF [40]. Six studies explored *participants' engagement* through their *satisfaction* of the intervention [34,36,37,39,42,45]. The *quality of delivery* was mentioned in three articles [38–40], but was considered as an indicator of IF rather than as a moderator. A number of articles also provided information about the *recruitment of participants*. These included patients from the center where the study took place [33,34,43,45], a subsample of a national health survey [35], people with diabetes living close to the study center [27], as well as persons recruited through mailing and advertisements [36,38,39], through a regional practice-based research network [37], a specific insurance plan [40], letters of invitation [41], and via screening [42,46]. One article

Table 2 Conceptualization and assessment of IF for empirical studies on self-management programs for people with or at risk of type II diabetes.

	Conceptualization of IF				Moderators of IF				Assessment of IF							
	Content	Frequency	Duration	Coverage	Intervention complexity	Facilitating strategies	Quality of delivery	Participant responsiveness	Recruitment	Context	Direct observation	Audio/video tap	Provider questionnaire or checklist	Provider interview	Participants questionnaire	Participants interview
Di Loreto et al. [33]	✓								✓							
Perrin et al. [34]	✓				✓	✓	✓	✓	✓	✓			✓			
Taub [35]	✓			✓	✓	✓	✓	✓	✓	✓				✓		
Huizinga et al. [27]	✓			✓	✓	✓	✓	✓	✓	✓						✓
Griffin et al. [36]	✓			✓	✓	✓	✓	✓	✓	✓						
Whittemore et al. [37]	✓			✓	✓	✓	✓	✓	✓	✓						
Castro et al. [38]	✓			✓	✓	✓	✓	✓	✓	✓						
Lakerveld et al. [39]	✓			✓	✓	✓	✓	✓	✓	✓						
Rothschild et al. [40]	✓			✓	✓	✓	✓	✓	✓	✓						
Siminerio et al. [41]	✓			✓	✓	✓	✓	✓	✓	✓						
Haltiwanger & Galindo [42]	✓			✓	✓	✓	✓	✓	✓	✓						
Schouten et al. [43]					✓	✓	✓	✓	✓	✓						
Cherrington et al. [44]					✓	✓	✓	✓	✓	✓						
Pagoto et al. [45]					✓	✓	✓	✓	✓	✓						
Kramer et al. [46]					✓	✓	✓	✓	✓	✓						
	11	4	3	3	0	14	6	14	3	3	3	5	1	8	1	

only provided information on the recruitment of the providers [44]. No conceptualization was found of the context or of the intervention complexity. However, some authors suggested in their discussion that the context could impact the IF [34,35].

4.2.3. Assessment of implementation fidelity

Ten studies report on the way in which IF was assessed [27,33–40], which includes observation as well as self-report measures completed by the providers or participants. Observation by researchers involved the recording of sessions [27,38,39] as well as direct observation [33,34,40]. Self-report assessment of IF by the participants was done through questionnaires or checklists [35,37–39,43] and interviews [34]. The same is found for self-report measures by the providers: 5 studies used questionnaires [36–40] and one study used interviews [34].

The assessment of IF was also considered for each component separately: The content of the education program was mainly assessed through observation [27,33,34,38–40] or by checklists completed by the providers [36–40]. Two studies used self-report measures (questionnaire or interviews) completed by the participants [34,35]. Overall, four studies combined observation and self-reported measures to assess the adherence to the program content. Two studies [41,42] provide no information on how the content was assessed. Frequency was mainly assessed through provider self-reports [36,37] and observation [33], whereas one article provides no information on the way frequency has been measured [42]. Duration was assessed through provider self-reports in the three studies [36–38], one of which combined this with an observational measure [38]. Coverage was assessed through self-reports completed by providers [36,37] or participants [35]. As for the moderators, participant responsiveness was primarily explored with self-report measures completed by the participants [34,37,39,42,45] or providers [36], and quality of delivery was assessed via observation [38–40]. Information about other moderators was reported but was not assessed through a specific measure.

4.2.4. Level of implementation fidelity

Eleven studies report results regarding the levels of IF that were attained [27,33–41,43]. Four of these studies give a global appreciation [33,34,40,41], while 7 express the results obtained for IF in percentages. The latter includes the percentage of sessions where the component was implemented [27,36,38]; the percentage of patients who received the intervention components [41,43]; the level of implementation of each component [39]; or the division of the number of components completed per session by the number of components of the protocol [37].

Three articles provide a global percentage for IF, which varies between 58.2% and 92%. This leads to the authors' overall conclusion that the interventions were delivered with a relatively high level of fidelity [27,36,37]. The percentage of IF that is reached also depends on the kind of component [27,35,36,38,43], with 94% of the patients receiving counseling to adopt a healthy diet, but only 52% receiving instructions to monitor blood glucose [43]. Two articles report an increase of IF as a result of monthly control reports [40] or provider education [41]. Another factor of influencing the percentage of IF is the person who does the coding [39]. Finally, one study concludes that providers do not give enough information to the patients in accordance to the ADA recommendations [35].

4.2.5. Implementation fidelity and program outcomes

Although it is generally assumed that the fidelity of implementation affects the results of the intervention, only one of the eleven articles in this review addresses the relationship between IF and intervention outcomes explicitly [36]. The paper describes two

DSM programs. For both programs, no significant relationship is found between the intervention dose across seven different sites and physical activity effect size. To explain this lack of effect, three possible factors are mentioned: the program may have been robust enough to not be affected by small to moderate adaptations; the variation in IF across sites may not have been sufficiently large to show a relationship with the outcomes; or the sample size across the sites may have been too small to detect a relationship.

5. Discussion and conclusion

5.1. Discussion

Despite the growing interest in the way health interventions are being implemented, only a very limited number of empirical studies have been published to date on efforts to conceptualize and assess IF and to study its impact on program outcomes. While this holds true for IF in general, it certainly also applies to DSM education. Although implementation is increasingly mentioned as an important topic in theoretical papers on DSM support, the notion of implementation remains a peripheral issue in existing systematic reviews of DSM education programs. Moreover, as appears from the review presented in this paper, few empirical studies of DSM education programs addressed the issue of IF, and when it is, the information provided is often incomplete. More specifically, information about the conceptualization and measurement of IF is often lacking, and the relationship between IF and program outcomes is hardly ever investigated. Despite this lack of information, some authors argue that they reached a high level of IF. This kind of affirmation may sometimes lead to confusion, as no information is provided about what IF then exactly entails.

Overall, the results of our review resemble those of literature reviews on IF with regard to interventions targeting other health conditions. Durlak and DuPre [26] concluded from a review of more than 500 studies of health promotion and prevention programs that the level of implementation in terms of variables related to communities, providers, the delivery system and the support system affects the outcomes, but that the collection of implementation data as an essential feature of program evaluation leaves much room for improvement. O'Donnell [47] reviewed 133 studies on IF and concluded that there are too few studies to guide researchers on how to measure the fidelity of the implementation of core curriculum interventions and relate the results to outcomes. This is particularly the case for efficacy and effectiveness studies, where the requirements for fidelity measures differ. Perepletchikova et al. [48], in a study of psychotherapy research, concluded that only 3.5% of the interventions adequately addressed IF.

The reasons why IF is seldom addressed in published intervention studies are manifold. For instance, researchers may acknowledge the importance of IF, but think they do not have enough knowledge to measure it [48]. Furthermore, journals do not require papers on intervention studies to report on IF, and the word limits may also be an obstacle to report on IF in the published literature. Finally, conducting an evaluation of IF is costly and time-consuming, and there are no agreed-upon procedures for its assessment.

Whereas theoretical models of IF have been developed and are increasingly used in the health education literature, the few articles that do provide information on IF in the context of DSM do not refer to these models to conceptualize the notion. Instead, they use isolated notions or related concepts to evaluate IF. The most commonly used indicator of IF in DSM education is adherence, evaluated through the correspondence of the program content with guidelines. This is again in keeping with reviews of IF in other areas. Durlak and DuPre [26] noted in their review of prevention

programs for children and adolescents that IF was mainly assessed through adherence to content and dose, and Dane and Schneider [21] concluded that researchers mainly focus on exposure and adherence to assess IF. Only a small number of studies on DSM education also consider frequency, coverage or duration to conceptualize adherence in terms of intervention dose, and a few studies focus on the quality of the delivery as an indicator of IF.

To measure IF of DSM programs, researchers resort to observation as well as self-report measures by either providers or participants, or both. Observation can yield information on specific aspects of the implementation (e.g. nonverbal interactions during the intervention), whereas self-report measures can provide more general information on aspects of the implementation [22]. However, neither one of these techniques is without problems. Direct observation requires much time, may influence the intervention delivery, and may not always be possible, while provider self-reports are likely to suffer from social desirability bias, and participant self-reports may be influenced by the participants' lack of familiarity with evaluation procedures or by their unwillingness to evaluate the provider unfavorably. Consequently, the results of different types of measures are not necessarily correlated [22,49].

Our review also shows that different aspects of IF are typically assessed through different kinds of measures, whereby the content of the intervention is usually assessed through observation, while the different components of the intervention dose (i.e., frequency, duration, and coverage) are more often assessed through provider self-reports, and participant responsiveness via self-report measures completed by the participants. However, the variety of methods that are used to assess the different components of IF is not necessarily the result of a well-informed, deliberated choice. According to Century et al. [18], researchers do not have a shared conceptual understanding of what fidelity of implementation is and how it can be measured, which means that the choice of measures is typically a secondary focus and based on specific contexts and programs.

Finally, it is also important to consider the potential barriers and moderators to IF. Although certain moderators were mentioned in the studies included in our review, none of these studies assessed their effects on IF. This finding is congruent with that of Appiah et al. [50], who also noted that factors affecting the implementation of DSM guidelines remain largely unstudied.

Measuring IF is not a goal in itself, but serves the purpose of improving the way programs are implemented and, thereby, enhancing their effectiveness. It seems reasonable to assume that IF can be significantly improved by using well-developed guidelines. Some authors have proposed guidelines to develop implementation protocols that pay explicit attention to IF and offer clear guidance to improve fidelity (e.g. [51]). However, the indicators to measure IF according to these guidelines are specific to the programs concerned, and can therefore not be used to compare different programs in terms of their IF.

5.2. Study limitations

While this paper is the first systematic review of IF applied to DSM education programs, the limitations of this review should be acknowledged. A first limitation concerns the small sample of the review. This small sample size may be due to the restriction of the search terms. However, while additional terms such as "intervention compliance" or "implementation quality" could have been used to complement "implementation fidelity", preliminary research showed that these other terms did not lead to the identification of additional relevant articles. On the other hand, the addition of search terms like "self-care" in addition to self-management could have led to the identification of more

interventions, but would have implied a different focus. Second, the limited number of articles that met all the inclusion criteria makes it difficult to draw far-reaching conclusions from the content analyses. This is further complicated by the moderate inter-reviewer agreement for the application of the content criteria, which again illustrates the difficulty of finding a common agreement regarding the conceptual definition of IF. On the other hand, our findings concur with those of others reporting that many implementation studies in the literature do not address sufficiently the IF [52]. Moreover, the fact that our conclusions concur with those from reviews on IF applied to other areas also attests to their validity.

Another limitation is that the articles included in this review may not fully represent the variety of existing DSM programs. For instance, no studies were found that used an IT-based intervention, which is an approach that is becoming increasingly popular on account of its benefits for patients with limited literacy and greater self-management support needs [53,54]. As the role of providers is different in IT-based intervention, it would be interesting to see how IF would be assessed in such interventions.

Finally, as IF is a relatively new concept, a unanimously agreed-upon framework with regard to this concept is not yet available. Amongst the various approaches that have been proposed, the model proposed by Carroll et al. [15] was chosen as a basis for this review, as this provides the most complete framework on IF currently available. As a result, all publications included in this review were content analyzed using his specific framework. It is possible that the use of another framework would have yielded somewhat different results and interpretations. However, this model has been selected for its completeness, and another approach to conceptualize IF, such as the critical components approach, would not have provided the possibility to identify the different dimensions assessed in this review.

5.3. Conclusion

Self-management education is currently a standard of medical care in diabetes [55]. In 2006, the American Association of Diabetes Educators, the American Diabetes Association, the American Dietetic Association, the Veteran's Health Administration, and other organizations jointly defined new DSM education standards [56]. Of the ten standards that were defined, three are related to facilitating strategies to improve the IF: documenting organizational structure, mission statement and goals in order to provide a solid basis to deliver quality diabetes education (standard 1); the presence of a designated coordinator to ensure that diabetes education is delivered through a coordinated and systematic process (standard 4); and regular continuing education for the providers (standard 5). This fact alone illustrates the growing importance of IF for the practice of DSM education. Despite this development, however, there is a dearth of empirical studies investigating IF for DSM programs using well-defined, theory-based and valid measures. Perhaps even more surprisingly, only one study assessed the relationships between IF components, moderators and program outcomes.

5.4. Practice implications

The results of this review suggest that reports on self-management education programs for diabetes patients should systematically describe how the program was implemented, how IF was conceptualized, what components and moderators were considered and how they were assessed, and how the IF affected the program outcomes. To facilitate this type of data collection and

reporting, we argue that it is essential to develop a shared understanding of what fidelity of implementation entails and how it can be measured. In addition to agreeing on a common conceptual framework to understand IF, it is essential to develop common, easy-to-use measures to assess the core components of IF across multiple programs. Apart from facilitating the monitoring of the implementation of a given education program, such tools would allow a comparison of different programs in terms of their IF.

Attempts to develop such tools have already been reported in the context of process evaluation of comprehensive health promotion interventions [18,51]. Although some authors are skeptical about the relevance of generic assessment tools for IF of DSM programs and argue that specific programs need specific fidelity measures [47], the exclusive use of program-specific measures to assess IF creates a barrier to IF assessment in practice. Most program providers and researchers do not have the time or resources to develop measures of IF themselves. As such, a generic, user-friendly tool based on a relevant conceptual model need not replace the level of detail that program-specific measures can provide, but could nevertheless enhance the capacities of program providers and researchers to assess core aspects of the quality with which DSM programs are implemented, and compare them to other programs.

I confirm that the patient/person(s) have read this manuscript and given their permission for it to be published in PEC.

Conflict of interest

The three authors have no conflict of interest.

Acknowledgements

This article is part of the Diabetes Literacy project supported by grant FP7-Health-2012-Innovation-1/306186 of the European Commission. The authors would like to thank the partners of the Diabetes Literacy Consortium for contributing to the content of this article.

Appendix A

The Diabetes Literacy Consortium consists of the following members:

Université Catholique de Louvain, Belgium: Stephan Van den Broucke, Gerard Van der Zanden, Marie Housiaux, Louise Schinckus.

Technische Universität Dresden, Medical Faculty, Germany: Peter Schwarz, Gabriele Mueller, Georg, Gallein, Thoralf Stange, Christina Ruch, Antje Linder.

University College Dublin, Ireland: Gerardine Doyle, Kenneth Cafferkey, Sarah Gibney, Emily Carroll.

Ludwig Boltzmann Institute for Health Promotion Research, Austria: Jürgen Pelikan, Florian Röthling, Kristin Ganahl.

Maastricht University, Department of International Health, The Netherlands: Helmut Brand, Kristine Sörensen.

University of Southampton, UK: Lucy Yardley, Ali Rowsell, Ingrid Muller.

CLALIT Health Services, Israel: Diane Levin, Ziv Har-Gil.

University of California at San Francisco, USA: Dean Schillinger, Courtney Lyles.

Taipei Medical University, Taiwan: Peter Chang.

References

- [1] International Diabetes Federation. *Diabetes atlas*. 5th ed. Brussels: International Diabetes Federation; 2012.
- [2] Health Service Executive. National service plan; 2011. Available from: <http://www.hse.ie/eng/services/Publications/corporate/nsp2011.pdf> [23.11.13].

- [3] Hjartåker A, Langseth H, Weiderpass E. Obesity and diabetes epidemics: cancer repercussions. *Adv Exp Med Biol* 2008;630:72–93.
- [4] Bartlett E. Historical glimpses of patient education in the United States. *Patient Educ Couns* 1986;8:135–49.
- [5] Fond-Harmant L. Diabetes plans in six European countries and in Canada. *Santé Publique* 2011;23:41–53.
- [6] Clement S. Diabetes self-management education. *Diab Care* 1995;18:1204–14.
- [7] Norris SL, Engelgau MM, Venkat-Narayan KM. Effectiveness of self-management training in type 2 diabetes: a systematic review of randomized controlled trials. *Diab Care* 2001;24:561–87.
- [8] American Association Diabetes Educators. 7TM self-care behaviors position statement 2011; 2013. Available from: http://www.diabeteseducator.org/export/sites/aade/_resources/pdf/research/AADE7_Position_Statement_2011.pdf [21.05.13].
- [9] Brown SA. Interventions to promote diabetes self-management: state of the science. *Diab Educ* 1999;25:52–61.
- [10] Duke SA. Individual patient education for people with type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2009;21:1–40.
- [11] Pal K, Eastwood S, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus (Review). *Cochrane Libr* 2013;3:1–150.
- [12] Simmons D, Cohn S, Bunn C, Birch K, Donald S, Paddison C, et al. Testing a peer support intervention for people with type 2 diabetes: a pilot for a randomised controlled trial. *BMC Fam Pract* 2013;14:5.
- [14] Engelgau MM, Venkat-Narayan KM, Saaddine JB, Vinicor F. Addressing the burden of diabetes in the 21st century: better care and primary prevention. *J Amer Soc Nephrol* 2003;14:88–91.
- [15] Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci* 2007;22:40.
- [16] Dusenbury L, Brannigan R, Falco M, Hansen WB. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res* 2003;18:237–56.
- [17] Dobson LD, Cook TJ. Avoiding Type III error in programme evaluation: results from a field experiment. *Eval Program Plan* 1980;3:269–76.
- [18] Century J, Rudnick M, Freeman C. A framework for measuring fidelity of implementation: a foundation for shared language and accumulation of knowledge. *Amer J Eval* 2010;31:199–218.
- [19] Hill LG, Owens RW. Component analysis of adherence in a family intervention. *Health Educ* 2013;113:264–80.
- [20] Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity criteria: development, measurement, and validation. *Amer J Eval* 2003;24:315–40.
- [21] Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control. *Clin Psychol Rev* 1998;18:23–45.
- [22] Schoenwald SK, Garland AF, Chapman JE, Frazier SL, Sheidow AJ, Southam-Gerow MA. Toward the effective and efficient measurement of implementation fidelity. *Adm Policy Ment Health* 2010;38:32–43.
- [23] Hasson H. Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci* 2010;5:67.
- [24] Horner S, Rew L, Torres R. Enhancing intervention fidelity: a means of strengthening study impact. *J Specialists Pediatr Nurs* 2006;11:80–9.
- [25] Santacroce SJ, Maccarelli LM, Grey M. Intervention fidelity. *Nurs Res* 2004;53:63–6.
- [26] Durlak JA, DuPre EP. Implementation matters a review of research on the influence of implementation on programme outcomes and the factors affecting implementation. *Amer J Comm Psychol* 2008;41:327–50.
- [27] Huizinga MM, Shintani A, Michon S, Brown A, Wolff K, Shackelford L, et al. A randomized controlled trial to prevent glycaemic relapse in longitudinal diabetes care: study protocol. *Implement Sci* 2006;1:24.
- [28] Glasgow RE. A practical model of diabetes management and education. *Diab Care* 1995;18:117–26.
- [29] Leeman J, Jackson B, Sandelowski M. An evaluation of how well research reports facilitate the use of findings in practice. *J Nurs Scholarship* 2006;38:171–7.
- [30] McAndrew LM, Musumeci-Szabó TJ, Mora PA, Vileikyte L, Burns E, Halm EA, et al. Using the common sense model to design interventions for the prevention and management of chronic illness threats: from description to process. *Brit J Health Psychol* 2008;13:195–204.
- [31] Hayes E, McCahon C, Panahi MR, Hamre T, Pohlman K. Alliance not compliance: coaching strategies to improve type 2 diabetes outcomes. *J Amer Acad Nurs Pract* 2008;20:155–62.
- [32] Hunt J. Motivational interviewing and people with diabetes. *Eur Diab Nurs* 2011;8:68–73.
- [33] Di Loreto C, Fanelli C, Lucidi P, Murdolo G, De Cicco A. Validation of a counseling strategy to promote the adoption and the maintenance of physical activity by type 2 diabetic subjects. *Diab Care* 2003;26:404–8.
- [34] Perrin M, Burke SG, O'Connor D, Walby G, Shippey C, Pitt S, et al. Factors contributing to intervention fidelity in a multi-site chronic disease self-management programme. *Implement Sci* 2006;1:26.
- [35] Taub L-FM. Concordance of provider recommendations with American Diabetes Association's guidelines. *J Amer Acad Nurs Pract* 2006;185:124–33.
- [36] Griffin SF, Wilcox S, Ory MG, Lattimore D, Leviton L, Castro C, et al. Results from the active for life process evaluation: programme delivery fidelity and adaptations. *Health Educ Res* 2010;25:325–42.
- [37] Whittemore R, Melkus GE, Alexander N, Zibel S, Visone E, Muench U, et al. Implementation of a lifestyle programme in primary care by nurse practitioners. *J Amer Acad Nurse Pract* 2010;22:684–93.
- [38] Castro CM, Pruitt LA, Buman MP, King AC. Physical activity programme delivery by professionals versus volunteers: the TEAM randomized trial. *Health Psychol* 2011;30:285–94.
- [39] Lakerveld J, Bot S, Chinapaw M, Tulder Mv Kingo L, Nijpels G. Process evaluation of a lifestyle intervention to prevent diabetes and cardiovascular diseases in primary care. *Health Promot Pract* 2012;13:696–706.
- [40] Rothschild SK, Martin MA, Swider SM, Lynas CT, Avery EF, Janssen I, et al. The Mexican-American Trial of Community Health workers (MATCH): design and baseline characteristics of a randomized controlled trial testing a culturally tailored community diabetes self-management intervention. *Contemp Clin Trials* 2012;33:369–77.
- [41] Siminerio LM, Piatt G, Zgibor JC. Implementing the chronic care model for improvements in diabetes care and education in a rural primary care practice. *Diab Educ* 2005;31:225–34.
- [42] Haltiwanger EP, Galindo D. Reduction of depressive symptoms in an elderly Mexican-American female with type 2 diabetes mellitus: a single-subject study. *Occup Ther Intern* 2013;20:35–44.
- [43] Schouten LM, Hulscher ME, van Everdingen JJ, Huijsman R, Niessen L, Grol RP. Short- and long-term effects of a quality improvement collaborative on diabetes management. *Implement Sci* 2010;5:94.
- [44] Cherrington A, Ayala G, Amick H, Allison J, Corbie-Smith G, Scarinci I. Implementing the community health worker model within diabetes management: challenges and lessons learned from programmes across the United States. *Diab Educ* 2008;34:824–33.
- [45] Pagoto SL, Kantor L, Bodenlos JS, Gitkind M, Ma Y. Translating the diabetes prevention programme into a hospital-based weight loss programme. *Health Psychol* 2008;27:S91–8.
- [46] Kramer MK, Kriska AM, Venditti EM, Miller RG, Brooks MM, Burke LE, et al. Translating the diabetes prevention programme: a comprehensive model for prevention training and programme delivery. *Amer J Prev Med* 2009;37:505–11.
- [47] O'Donnell C. Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Rev Educ Res* 2008;78:33–84.
- [48] Perepletchikova F, Cheriji E, Hilt LM, Kazdin AE. Barriers to implementing treatment integrity procedures: survey of treatment outcome researchers. *J Consult Clin Psychol* 2009;77:212–8.
- [49] Ellis DA, Naar-King S, Templin T, Frey MA, Cunningham PB. Improving health outcomes among youth with poorly controlled type 1 diabetes: the role of treatment fidelity in a randomized clinical trial of multisystemic therapy. *J Fam Psychol* 2007;21:363–71.
- [50] Appiah B, Hong Y, Ory MG, Begaye D, Bolin J, Forjuoh S. Challenges and opportunities for implementing diabetes self-management guidelines. *J Amer Board Fam Med* 2013;26:90–2.
- [51] Saunders RP, Evans MH, Joshi P. Developing a process-evaluation plan for assessing health promotion programme implementation: a how-to guide. *Health Promot Pract* 2005;6:134–47.
- [52] Perepletchikova F, Treat AT, Kazdin AE. Treatment integrity in psychotherapy research: analysis of the studies and examination of associated factors. *J Consult Clin Psychol* 2007;75:829–41.
- [53] Schillinger D, Hammer H, Wang F, Palacios J, McLean I, Tang A, et al. Seeing in 3-D: examining the reach of diabetes self-management support strategies in a public health care system. *Health Educ Behav* 2007;35:664–82.
- [54] Schillinger D, Handley M, Wang F, Hammer H. Effects of self-management support on structure, process, and outcomes among vulnerable patients with diabetes. *Diab Care* 2009;32:559–66.
- [55] ADA. Standards of medical care in diabetes-2013. *Diab Care* 2013;36: S11–66.
- [56] Funnell MM, Brown TL, Childs BP, Haas LB, Hoseney GM, Jensen B, et al. National standards for diabetes self-management education. *Diab Care* 2010;33: 589–96.