

# Meta-analysis does not allow appraisal of complex interventions in diabetes and hypertension self-management: a methodological review

M. Lenz · A. Steckelberg · B. Richter · I. Mühlhauser

Received: 15 January 2007 / Accepted: 20 February 2007 / Published online: 23 May 2007  
© Springer-Verlag 2007

**Abstract** Common methodologies used in systematic reviews do not allow adequate appraisal of complex interventions. The aim of the present study was to describe and critically appraise current methods of systematic reviews on complex interventions, using diabetes and hypertension patient education as examples. PubMed, the Cumulative Index to Nursing and Allied Health (CINAHL), the Cochrane Library and Health Technology Assessment databases were searched. Systematic reviews focusing on diabetes or hypertension patient education were included. Authors were contacted. Two investigators independently evaluated the reviews. The available evidence of three patient education programmes of diabetes and hypertension self-management implemented in Germany was used as a reference. We included 14 reviews; 12 reviews mentioned that the included education programmes were multidimensional. Reviews on comparable topics identified different publications of the same programme. Identical programmes were classified differently within and between reviews.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00125-007-0679-z) contains supplementary material, which is available to authorised users.

M. Lenz (✉) · A. Steckelberg  
Unit of Health Sciences and Education,  
University of Hamburg,  
Martin-Luther-King Platz 6,  
20146 Hamburg, Germany  
e-mail: matthias.lenz@uni-hamburg.de

B. Richter  
Department of General Practice, University Hospital,  
Heinrich-Heine University,  
Düsseldorf, Germany

I. Mühlhauser  
Unit of Health Sciences and Education, University of Hamburg,  
Hamburg, Germany

Education programmes were dissected to analyse effects of single components. Different components of identical programmes were used. Interdependencies between components were not explored. Six reviews performed meta-analysis across programmes with heterogeneous educational or organisational approaches. The complexity of efficacy measures was disregarded: e.g. HbA<sub>1c</sub> was used as an isolated outcome variable without considering treatment goals, effects on hypoglycaemia, body weight or quality of life. Our results indicate that methods of current systematic reviews are not fully equipped to appraise patient education and self-management programmes. Since these are complex and heterogeneous interventions, consideration of aggregated evidence is necessary.

**Keywords** Diabetes · Education · Hypertension · Methods · Self-management · Systematic reviews

## Abbreviations

RCT randomised controlled trial  
UKMRC UK Medical Research Council

## Introduction

In diabetes care there are single interventions (e.g. a single drug) and complex interventions (e.g. treatment regimens or diabetes self-management). The latter comprise a number of separate elements (active components), all of which seem essential to their proper functioning [1, 2]. A self-management programme for type 1 diabetes may illustrate the complexity of such interventions [3]. Decisive components are the insulin regimen used and the quality of the teaching process to empower patients to carry it out effectively and safely. Empowerment of patients to set individual treatment

goals and to balance favourable blood glucose targets and an acceptable risk of hypoglycaemia by self-adaptation of insulin dosages to adjust to lifestyle may be more effective than defining normoglycaemia as the primary treatment goal and asking patients to adapt their lifestyle to match prescribed doses of insulin [4]. Liberalisation of the diet may be important to motivate patients to carry out an intensified insulin therapy regimen in the long term [5]. Though indispensable, knowledge by itself may not improve outcomes. The information and how it is transmitted are decisive [3]. Blood glucose self-monitoring may be at best useless unless patients have learned to interpret results and to react by adjusting insulin dosages [3].

A high-quality randomised controlled trial (RCT) is considered the most valid method of evaluating a medical intervention, and a systematic review of high-quality RCTs the most powerful evidence available [6]. A systematic review is a summary of the medical literature that uses explicit methods to systematically search, critically appraise and synthesise the literature on a specific issue [6]. They may, but need not, include a meta-analysis as a statistical method for combining the results of individual studies [6].

The UK Medical Research Council (UKMRC) has proposed a framework that allows a systematic and transparent evaluation of complex interventions. Five sequential phases of a ‘continuum of increasing evidence of complex interventions’ (hereafter referred to as ‘increasing evidence’) have been defined, which require both qualitative and quantitative evidence [1, 2]

#### Phases of ‘increasing evidence of complex interventions’ [1, 2]

- The theoretical or preclinical phase: to explore the relevant theory and to identify evidence showing that the complex intervention may work.
- Phase 1—the modelling phase: to delineate the components of the complex intervention and the underlying mechanisms involved.
- Phase 2—the phase of exploratory trials: to test the feasibility of delivering the complex intervention and its acceptability to providers and patients.
- Phase 3—the main efficacy trial: controlled trial(s) to compare the complex intervention with an appropriate alternative.
- Phase 4—the phase of long-term implementation: to establish the transferability and reproducibility and the long-term and real-life effectiveness of the complex intervention.

It has been suggested that, for many reasons, current methods used in systematic reviews do not allow adequate appraisal of complex interventions such as diabetes or

hypertension self-management programmes [7]. For example, systematic reviews do not consider the theory behind the complex intervention (e.g. behavioural models) and do not differentiate between trials designed to determine efficacy and those focusing on implementation.

If used appropriately, meta-analysis is a powerful tool for investigating overall effects. However, if studies are clinically or methodologically heterogeneous, data-pooling may be meaningless and genuine differences in effects may be obscured [8]. Complex interventions are heterogeneous in their goals, methods and target populations [9–14]. Thus, using meta-analysis to evaluate complex interventions may disregard the complexity of the efficacy measures used in the original studies. For example, HbA<sub>1c</sub> is used as a single outcome variable without considering individual treatment goals or effects on hypoglycaemia, body weight or quality of life [9, 15, 16]. An increase in HbA<sub>1c</sub> from 6.5 to 7% might be considered a deterioration, but would clearly be a desirable outcome if it halved the rate of hypoglycaemia.

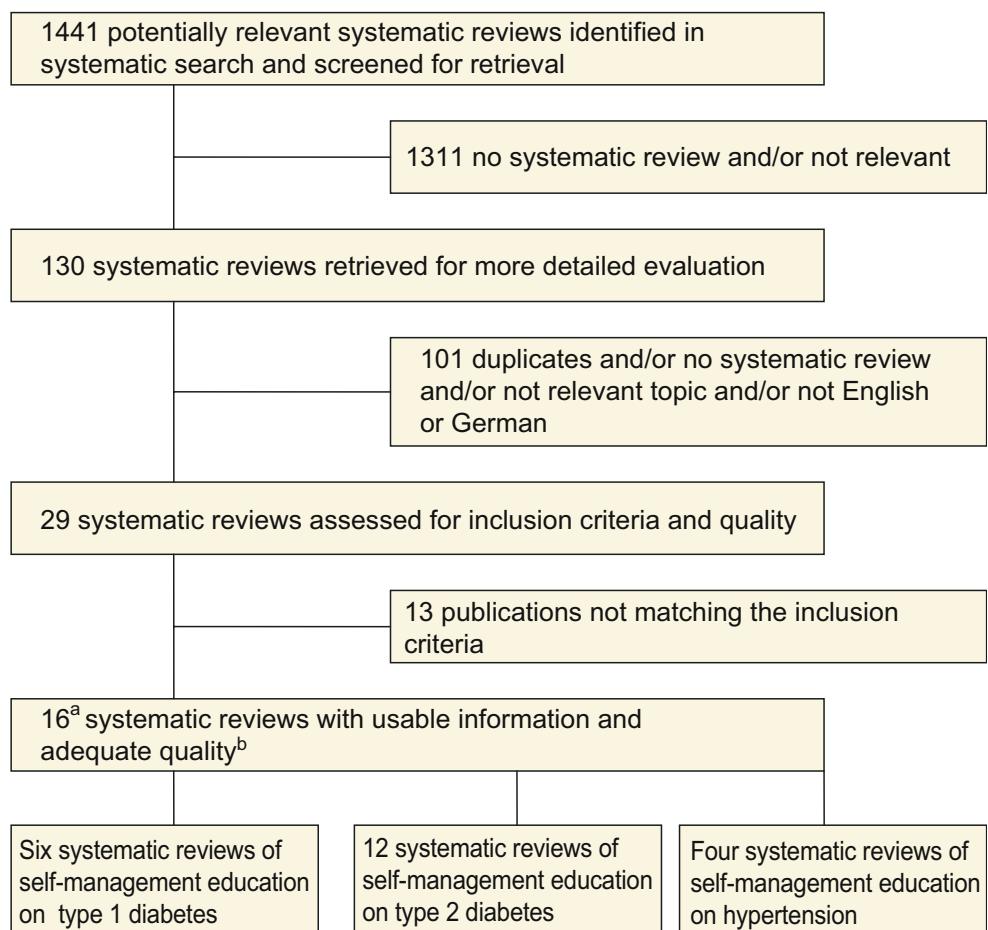
The aim of the present study was to describe and critically appraise available methodologies of systematic reviews on complex interventions. Patient self-management education programmes of diabetes and hypertension implemented in Germany were used as examples.

#### Materials and methods

The Cochrane Library, PubMed, Cumulative Index to Nursing and Allied Health (CINAHL) and Health Technology Assessment databases were systematically searched for systematic reviews published between 1997 and March 2006 (see search strategy in [Electronic supplementary material \[ESM\]](#)). Reference lists of retrieved reviews were screened for additional publications. Reviews were included if patient education programmes of diabetes and hypertension self-management were evaluated (see Selection criteria in [ESM](#)). Included reviews were analysed independently by two investigators (M. Lenz and A. Steckelberg) using standardised data extraction forms created following Cochrane criteria [17], according to predefined hypotheses (see Fig. 1, and Hypotheses and Data synthesis in [ESM](#)). All authors of the included reviews were contacted after data synthesis and a preliminary version of our review was provided. They were asked whether they felt that their review had been considered appropriately and their findings interpreted correctly.

The ‘increasing evidence’ of three patient self-management education programmes—type 1 diabetes, type 2 diabetes and hypertension—was used as the reference. It was assessed: (1) whether the selection criteria and search strategies used in the included systematic reviews were appropriate for detection of the ‘increasing evidence’ of

**Fig. 1** Flow diagram of study selection. The term ‘systematic review’ includes meta-analyses and health technology assessment reports. Fourteen reviews in 16 publications (*superscript a*). Some reviews address more than one topic (*superscript b*)



the three reference programmes (ESM Tables 1, 2, 3); (2) which publications related to our three reference programmes had been identified; (3) whether theoretical background literature (theoretical/preclinical phase) had been considered; (4) which active components had been identified and included (ESM Tables 4, 5, 6); (5) whether classification of interventions matched the type of intervention; and (6) whether all relevant patient outcomes had been included (ESM Tables 7, 8, 9).

## Results

A total of 15 reviews (16 publications) were included (Table 1, excluded reviews are listed in ESM Table 10). Various types of interventions were investigated (e.g. individual and group counselling, and structured training courses for patients). A meta-analysis was performed in eight reviews [9–11, 13, 15, 16, 18–20]. A total of six reviews exclusively used qualitative data synthesis [21–27] (ESM Table 11).

The majority of search strategies were comprehensive and transparently reported (ESM Table 12). In eight

reviews, experts in the field were contacted to identify additional publications [10, 11, 13, 15, 18–20, 25]. Reference tracking was performed in 11 reviews [10, 11, 13, 15, 16, 18–22, 24]. Predefined contact of all authors was not reported in any review.

In six reviews only RCTs were considered [9, 10, 13, 15, 18, 26, 27] (see ESM Table 12). Following the selection criteria of the reviews, 16 publications [28–43] referring to our three reference programmes should have been identified. In practice, a total of 11 were identified [5, 28, 30, 32, 34, 38, 39, 43–46], and seven [28, 32, 34, 35, 38, 39, 46] were included in at least one of the reviews. In six reviews the main controlled trials of our reference programmes [28, 30–32, 38, 44] were included in data synthesis [10, 11, 16, 18, 19, 25]. Out of 25 [5, 33–37, 39, 41–43, 45–59] replication trials (phase 4, see text box: Phases of ‘increasing evidence of complex interventions’) four [34, 35, 39, 46] were included in six reviews [16, 19, 21, 22, 24, 25]. Different reviews on identical topics included different publications referring to the same reference programme.

No review explicitly evaluated publications on the theoretical basis of the reference programmes (preclinical/theoretical phase of ‘increasing evidence’). The UKMRC

**Table 1** Characteristics of included reviews

| First author   | Topic   | Number of included studies  | Type of data synthesis   |
|----------------|---|---|--|
| Boulware [11]  | Patient-centred counselling, patient self-monitoring of BP and structured training courses on hypertension  | n=16  | 1. Meta-analysis: overall summary effects across pre-specified categories <sup>b</sup> of interventions according to study characteristics:<br>Single-intervention analysis (n=12): the intervention was part of a study comparing more than one intervention, but the remaining interventions of the study not meeting the inclusion criteria were excluded from analysis<br>Between-intervention analysis (n=7): the intervention group was compared with a control group also containing an intervention meeting the eligibility criteria or usual care<br>2. Meta-analysis: subgroup-analysis according to active components; articles (n=5) focusing on counselling by categorising articles by publication year; leader of intervention; duration of intervention; age of participants; percentage of white participants; and quality scores of articles   |
| Corabian [24]  | Outpatient education for self-management in adults with type 2 diabetes   | n=24 <sup>a</sup>   | 1. Qualitative assessment of quantitative studies within categories according to study characteristics; meta-analyses; primary studies on effectiveness; comprehensive and systematic reviews; Canadian studies<br>2. Qualitative assessment of qualitative research to identify factors that potentially influence the impact of the diabetes care and education regimen  |
| Deakin [19]    | Group-based, patient-centred training on type 2 diabetes  | n=11 (in 14 publications)   | 1. Meta-analysis: overall summary effects across pre-specified categories of interventions according to study characteristics <sup>c</sup> ; short-term follow-up (4–6 months); medium-term follow-up (6–12 months); long-term follow-up ( $\geq$ 12 months)<br>2. Descriptive summary according to outcome measures <sup>d</sup><br>3. Meta-analysis: overall summary effects across pre-specified category: follow-up term (12–14 months)<br>4. Meta-analysis: overall summary effects across pre-specified category: follow-up term 2 years<br>5. Meta-analysis: subgroup analysis according to ethnicity across specified categories of interventions: short-term follow-up (4–6 months), long-term follow-up (12–14 months)<br>6. Meta-analysis: subgroup analysis according to theoretical model; summary effect size calculated only across two studies, excluding those studies causing heterogeneity (only short-term follow-up [4–6 months])<br>7. Qualitative subgroup analysis according to the theoretical model <sup>d</sup><br>8. Meta-analysis: subgroup analysis according to the educator (nurse, dietitian or a combination of both); summary effect size calculated only across five studies, excluding those studies causing heterogeneity (only long-term follow-up [12–14 months])<br>9. Meta-analysis: subgroup analysis according to primary care interventions across specified categories <sup>c</sup> : short-term follow-up (4–6 months), long-term follow-up (12–14 months)<br>10. Qualitative subgroup analysis according to primary and secondary care <sup>d</sup><br>11. Qualitative subgroup analysis according to the number of participants in the group education programme for larger groups (16–18 participants) |
| Ellis [9]      | Patient education on type 1 and type 2 diabetes   | n=21 (in total); n=4 type 1 diabetes, n=14 type 2 diabetes, n=2 both, n=1 data not available n=59 [10]; n=56 [18] | 1. Meta-analysis: overall summary effects<br>2. Meta-analysis: subgroup analysis in categories of interventions according to active components: type of delivery; content; teaching method; intensity of intervention  |
| Fahey [10, 18] | Different strategies of care (e.g. self-management and patient education) used to improve the control and follow-up of patients with hypertension |   | Meta-analysis in pre-specified categories according to the types of intervention: self-monitoring education; education intervention directed at patients; education intervention directed at health professionals; health professional lead care; organisational interventions; appointment reminder systems   |

|                           |   |  |   |
|---------------------------|---|--|---|
| Gary [13]                 | Educational and behavioural interventions for individuals with type 2 diabetes  | <i>n</i> =18   | <ul style="list-style-type: none"> <li>1. Studies are descriptively summarised according to study characteristics and outcome measures</li> <li>2. Meta-analysis: overall summary effects</li> <li>3. Meta-analysis: subgroup analysis in categories of interventions according to active components: interventionist (physicians, nurse, dietician); mode of instruction (individual, group); topic of instruction (diet, medication, exercise, blood glucose self-monitoring)</li> </ul> <p>Studies are descriptively summarised according to outcome measures</p> <p>Studies are descriptively summarised within categories according to types of intervention: patient education for self-management in type 1 diabetes; patient education for self-management in type 2 diabetes; patient education for self-management in type 1 or type 2 diabetes</p>         |
| Jack [23]<br>Loveman [25] | Educational methods in type 2 diabetes<br>Structured educational interventions for diabetes self-management                                       | <i>n</i> =8<br><i>n</i> =24 (in total); <i>n</i> =4 (type 1 diabetes); <i>n</i> =16 (type 2 diabetes); <i>n</i> =4 (type 1 or type 2 diabetes) | <p><i>n</i>=63 (in total); <i>n</i>=21 (type 2 diabetes); <i>n</i>=24 (arthritis); <i>n</i>=18 (asthma)</p> <ul style="list-style-type: none"> <li>1. Studies are descriptively summarised within categories of interventions according to types of intervention: group-based; problem solving; stress management; cognitive behavioural; expressive writing; programme delivered by mail; combination of education with an action plan; exercise</li> <li>2. Qualitative analysis to examine whether particular outcomes were focused upon, those in each study were classified into seven broad categories</li> </ul> <p>Studies are descriptively summarised within categories of primary educational focus: knowledge or information (didactic) education or collaborative education); lifestyle behaviours; skill development; coping skills</p>                 |
| Newman [26]               | Self-management interventions for chronic illness (type 2 diabetes, arthritis and asthma)   | <i>n</i> =72   | <p><i>n</i>=31</p> <ul style="list-style-type: none"> <li>1. Meta-analysis: subgroup analysis according to study characteristics by the length of follow-up (follow-up <math>\geq</math>4 months after end of the intervention; follow-up 1–3 months after end of the intervention; follow-up during or immediately after the intervention)</li> <li>2. Meta-analysis: regression analysis to investigate potential treatment interactions: patient age; baseline GHb; treatment (insulin, diet only, oral hypoglycaemic agents); number of contacts; total contact time; time frame over which the intervention was delivered; group vs individual presentation of the intervention; who delivered the intervention; educational focus (lifestyle, skills, knowledge, coping skills or mixed); follow-up in months; setting in the USA vs other countries</li> </ul> |
| Norris [27]               | Self-management training in type 2 diabetes   | <i>n</i> =31   | <p><i>n</i>=30 (in total); <i>n</i>=8 (community gathering places); <i>n</i>=10 (home); <i>n</i>=10 (summer camps); <i>n</i>=1 (schools); <i>n</i>=1 (worksite)</p> <ul style="list-style-type: none"> <li>1. Meta-analysis: overall summary effect sizes for evidence on effectiveness, summarised for each setting (community gathering places, home, camps, schools and worksite)<sup>a</sup></li> <li>2. Qualitative analysis: assessment within categories, summarised for each setting (community gathering places, home, camps, schools, and worksite): applicability; other positive and negative effects, economic efficiency; barriers for implementation</li> <li>3. Additional descriptive information on positive and negative effects, applicability, economic efficiency, and barriers for implementation are described for each study</li> </ul>      |
| Norris [15]               | Self-management training in type 1 and type 2 diabetes in community settings  | <i>n</i> =41   | <p><i>n</i>=30 (in total); <i>n</i>=8 (community gathering places); <i>n</i>=10 (home); <i>n</i>=10 (summer camps); <i>n</i>=1 (schools); <i>n</i>=1 (worksite)</p> <ul style="list-style-type: none"> <li>1. Qualitative analysis: assessment whether the effectiveness of the interventions are influenced by pre-specified characteristics: types of intervention (professional, organisational, or financial); source of information (e.g. whether the intervention was carried or supported by professional organisation); type of diabetes</li> <li>2. Additional descriptive information on effects are described for each study categorised according to types of intervention:</li> </ul> <p>Professional<br/>Organisational<br/>Professional and organisational in combination</p>  |
| Renders [22]              | Interventions targeted at health professionals and/or the structure of care to improve the management of patients with type 1 and type 2 diabetes |  | <p>3. Qualitative analysis: assessment whether the effectiveness of the interventions are influenced by active components of the programmes:<br/>Professional (professional interventions combined with patient education; postgraduate education of</p>  |

**Table 1** (continued)

| First author | Topic   | Number of included studies | Type of data synthesis  |
|--------------|---|----------------------------|---|
| Renders [21] | Interventions targeted at health professionals and/or the structure of care to improve the management of patients with type 1 and type 2 diabetes | n=41                       | <p>healthcare professionals combined with local consensus procedures and/or audit and feedback; education for both healthcare professionals and patients)</p> <p>Organisational (patient education and learner-centred counselling; nurse or pharmacist assumed part of the physician's role and provided diabetes care in combination with a patient-oriented intervention; multidisciplinary teams in combination with arrangements for follow-up and patient education)</p> <p>Professional and organisational in combination (healthcare professionals received education through educational materials, educational meeting or both; arranging follow-up, audit and feedback, or generating reminders, or a combination; including patient education; involvement of nurses in diabetes management; including a telecommunication system)</p> <p>1. Qualitative analysis: effects are described for each study, categorised according to types of intervention:</p> <ul style="list-style-type: none"> <li>Professional</li> <li>Organisational</li> </ul> <p>2. Qualitative analysis: assessment whether the effectiveness of the interventions are influenced by active components of the programmes</p> <p>Professional (professional interventions combined with patient education; postgraduate education of healthcare professionals combined with local consensus procedures and/or audit and feedback; education for both healthcare professionals and patients)</p> <p>Organisational (patient education and learner-centred counselling (nurse or pharmacist assumed part of the physician's role and provided diabetes care in combination with a patient-oriented intervention; multidisciplinary teams in combination with arrangements for follow-up and patient education)</p> <p>Professional and organisational in combination (healthcare professionals received education through educational materials, educational meeting or both; arranging follow-up, audit and feedback, or generating reminders, or a combination; including patient education; involvement of nurses in diabetes management; including a telecommunication system)</p> <p>1. Meta-analysis: overall summary effect sizes only for evidence on efficacy for each outcome</p> <p>2. Meta-analysis: regression across chronic diseases to identify which components were associated with greater clinical benefits. The dependent variable was the summary effect size across all diseases. Indicator variables were each disease and its endpoint. Independent variables were percentage of dropouts, number of educational sessions, programme duration, programme format, education mode (face-to-face education), and reference to a behavioural science model</p> |
| Warsi [16]   | Patient self-management educational programmes for chronic diseases (e.g. diabetes and hypertension)  | n=71                       | <p>GHB glycated haemoglobin (HbA<sub>1</sub>, HbA<sub>1c</sub>)</p>   |

<sup>a</sup>The authors summarised seven studies to point out the research that had taken place over the last 10 years in Canada, although they did not meet their inclusion criteria; <sup>b</sup> some interventions (n=3) are allocated into both categories; <sup>c</sup> studies were excluded from analysis if heterogeneity was substantial; <sup>d</sup> if there was substantial heterogeneity between studies meta-analysis was not performed; <sup>e</sup> the effectiveness of interventions for type 1 and type 2 diabetes was examined separately

approach of the evaluation of complex interventions [1, 2] was cited in three reviews [18, 19, 25]. We did not find that the approach influenced their methodology.

All included reviews reported at least some features of the assessed interventions, which were either active components (e.g. setting, duration, interventionist, formal syllabus) or study characteristics (e.g. follow-up, age of participants, study quality). The investigated features were heterogeneous. Six reviews used regression or subgroup analysis to analyse the effect of single active components on outcomes [9, 11, 13, 15, 16, 19]. Descriptive analysis of the influence of active components was performed in five reviews [19, 21, 22, 24, 26]. Most of the components of our three reference programmes had been identified, but only a few had been evaluated. The identified and included components were different in reviews on identical topics (ESM Table 13).

In 12 reviews the included programmes were categorised by intervention [9–11, 13, 15, 16, 18, 19, 21, 22, 25–27]. Categories were defined according to the active components of the included interventions (e.g. setting, duration, interventionist, formal syllabus) [9, 11, 13, 15, 16, 19, 21, 22], the type of intervention (e.g. type of disease, type of activity, organisational interventions) [10, 18, 21, 22, 25–27] aimed at patients or health professionals, or the study characteristics (e.g. follow-up, age of participants, study quality) [11, 13, 15, 19, 24]. The applied categories varied.

The calculation of overall effects within categories of interventions was performed in eight reviews [9–11, 13, 15, 16, 18–20]. Nine reviews included qualitative data synthesis within categories [13, 19–22, 24–27] (Table 1). In some reviews [10, 11, 18] the allocation into categories did not meet the type of intervention [32, 34, 35].

No review included all relevant patient outcome measures of our reference studies (ESM Table 14). If quantitative reviews were aimed at exploring multiple outcome measures, multiple meta-analyses were performed [10, 11, 13, 16, 18–20]. If quantitative analysis was used according to the effect of active components, regression analysis or subgroup analysis was conducted [9, 11, 13, 15, 16, 19]. Interdependencies between different outcome measures of the same educational intervention were considered in only three reviews [25–27], all of which performed qualitative data synthesis.

Four out of 12 contacted authors responded (P. Corabian, Alberta Heritage Foundation for Medical Research, Edmonton, AB, Canada; T. Deakin, Pendle Rossendale Primary Care Trust, Burnley General Hospital, Burnley, UK; T. Fahey, University of Dundee, Dundee, UK; C. M. Renders, Institute of Research in Extramural Medicine, Amsterdam, The Netherlands). All of them replied that their reviews were, in general, considered appropriately and their findings interpreted correctly. Specific comments on methodological problems such as study selection and categorisation of components of education

programmes were acknowledged. We corrected five misinterpretations in reporting data and added unpublished information according to the authors' comments.

## Discussion

Most of the included systematic reviews discussed methodological challenges to appraise complex interventions. However, these considerations did not have adequate impact on the methods used in data synthesis.

The selection criteria used in most of the analysed reviews excluded study types other than RCTs; other important types of publications concerning the ‘increasing evidence’ were rarely included. The investigated reviews did not differentiate between the main controlled trial(s) and controlled replication trials referring to the same programme.

The importance of considering the theoretical basis of an education programme was widely discussed. However, we could not identify any approach to systematically assess the theoretical basis and its influence on judging the quality of the programmes.

The majority of reviews reported that the included programmes had been ‘multifaceted’ or ‘multidimensional’ or ‘consist of multiple active components.’ However, neither the UKMRC approach of ‘increasing evidence’ (published in 2000) [1, 2], nor a similar approach was integrated in any review. Methodological changes may take a long time to become accepted.

The categorisation of interventions used in the included reviews often seemed to be arbitrary: each review used different categories; none of them explained the rationale of their categorisation. Allocation of complex interventions into categories can be problematic, even if categories are derived from core components of programmes (e.g. education directed to the patient). If the categories refer to single but interdependent components, the compartmentalisation of efficacy is not possible.

Regression and subgroup analyses are the best tools for exploring heterogeneity [8]. However, these techniques should not be misused to identify the contributions of the various active components (e.g. intensity or duration of the programme) on the overall effect (e.g. knowledge of the target group or the importance perceived by the provider).

The analysed reviews did not consider all relevant patient outcome parameters. Components of complex outcome measures were singled out, especially if they used meta-analysis. The complex interdependency of individual treatment goals and outcomes (e.g. changes in medication and metabolic or blood pressure control) remained unexplored. Clinicians with experience in diabetes care might be consulted to focus the relevant clinical issues.

Recapitulating the research findings, we propose to take the following criteria into account when undertaking systematic reviews of complex interventions:

1. All studies referring to development, evaluation and implementation ('increasing evidence') should be considered, differentiating between phases of 'increasing evidence'.
2. Literature searches should therefore not be limited by criteria such as certain types of studies, specific target groups and publication dates; reference tracking should be performed and authors should be contacted systematically.
3. Active components should be described and assessed, but should only be examined separately if they are independent and should not be disassembled if they work interdependently.
4. Education programmes should not be allocated into categories referring to interdependent components.
5. All relevant patient-orientated outcome parameters should be included.
6. Pooling of outcome measures across different programmes is usually inappropriate. Instead, the relative importance of outcomes [60] and the complex interdependency between treatment goals and outcomes should be described in detail.

Information necessary for the evaluation of education programmes is difficult or impossible to identify. Therefore, specific search strategies need to be developed and validated that aim to identify publications on all phases of 'increasing evidence'. To facilitate the implementation of such an approach, what is needed is an electronic database that provides available programmes together with all relevant background information, or the implementation of such a system into existing databases.

**Acknowledgement** We thank G. Meyer (University of Bremen, Bremen, Germany) for her insightful comments on the draft version of this review.

**Duality of interest** None of the authors have a duality of interest with regard to this work.

## References

1. Health Services and Public Health Research Board (2000) A framework for development and evaluation of RCTs for complex interventions to improve health. UK Medical Research Council. <http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC003372>, last accessed in March 2007
2. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL et al (2000) Framework for design and evaluation of complex interventions to improve health. *BMJ* 321:694–696
3. Mühlhauser I, Berger M (1993) Diabetes education and insulin therapy: when will they ever learn? *J Intern Med* 233:321–326
4. Sämann A, Mühlhauser I, Bender R, Kloos C, Müller UA (2005) Glycaemic control and severe hypoglycaemia following training in flexible, intensive insulin therapy to enable dietary freedom in people with type 1 diabetes: a prospective implementation study. *Diabetologia* 48:1965–1970
5. Mühlhauser I, Bott U, Overmann H, Wagener W, Bender R, Jörgens V et al (1995) Liberalized diet in patients with type 1 diabetes. *J Intern Med* 237:591–597
6. Sackett DL, Strauss S, Richardson WS, Rosenberg W, Haynes RB (2000) Evidence-based medicine: how to practice and teach EBM, 2nd edn. Churchill Livingstone, Edinburgh
7. Mühlhauser I, Berger M (2002) Patient education-evaluation of a complex intervention. *Diabetologia* 45:1723–1733
8. Higgins JPT, Green S (2006) Collecting data. In: Cochrane handbook for systematic reviews of interventions 4.2.6. Wiley, Chichester, pp 99–137
9. Ellis SE, Speroff T, Dittus RS, Brown A, Pichert JW, Elasy TA (2004) Diabetes patient education: a meta-analysis and meta-regression. *Patient Educ Couns* 52:97–105
10. Fahey T, Schroeder K, Ebrahim S (2005) Interventions used to improve control of blood pressure in patients with hypertension. *Cochrane Database Syst Rev* 1:CD005182
11. Boulware LE, Daumit GL, Frick KD, Minkovitz CS, Lawrence RS, Powe NR (2001) An evidence-based review of patient-centered behavioral interventions for hypertension. *Am J Prev Med* 21:221–232
12. Brown SA (1992) Meta-analysis of diabetes patient education research: variations in intervention effects across studies. *Res Nurs Health* 15:409–419
13. Gary TL, Genkinger JM, Guallar E, Peyrot M, Brancati FL (2003) Meta-analysis of randomized educational and behavioural interventions in type 2 diabetes. *Diabetes Educ* 29:488–501
14. Greenhalgh T (1998) Meta-analysis is a blunt and potentially misleading instrument for analysing models of service delivery. *BMJ* 317:395–396
15. Norris SL, Lau J, Smith SJ, Schmid CH, Engelgau MM (2002) Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. *Diabetes Care* 25:1159–1171
16. Warsi A, Wang PS, LaValley MP, Avorn J, Solomon DH (2004) Self-management education programs in chronic disease: a systematic review and methodological critique of the literature. *Arch Intern Med* 164:1641–1649
17. Higgins JPT, Green S (2006) Collecting data. In: Cochrane handbook for systematic reviews of interventions 4.2.6. Wiley, Chichester, pp 91–96
18. Fahey T, Schroeder K, Ebrahim S (2005) Educational and organisational interventions used to improve the management of hypertension in primary care: a systematic review. *Br J Gen Pract* 55:875–882
19. Deakin T, McShane CE, Cade JE, Williams RD (2005) Group based training for self-management strategies in people with type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2:CD003417
20. Norris SL, Nichols PJ, Caspersen CJ et al (2002) Increasing diabetes self-management education in community settings. A systematic review. *Am J Prev Med* 22(Suppl 4):39–66
21. Renders CM, Valk GD, Griffin SJ, Wagner EH, Eijk Van JT, Assendelft WJ (2001) Interventions to improve the management of diabetes in primary care, outpatient, and community settings: a systematic review. *Diabetes Care* 24:1821–1833
22. Renders CM, Valk GD, Griffin S, Wagner EH, Eijk Van JT, Assendelft WJJ (2005) Interventions to improve the management of diabetes mellitus in primary care, outpatient and community settings. *Cochrane Database Syst Rev* 4:CD001481
23. Jack JL (2003) Diabetes self-management education research: an international review of intervention methods, theories, community partnerships and outcomes. *Dis Manag Health Outcomes* 7:415–428

24. Corabian P, Harstall C (2001) Patient diabetes education in the management of adult type 2 diabetes. Alberta Heritage Foundation for Medical Research HTA 23, Series A. AHFMR, Edmonton
25. Loveman E, Cave C, Green C, Royle P, Dunn N, Waugh N (2003) The clinical and cost-effectiveness of patient education models for diabetes: a systematic review and economic evaluation. *Health Technol Assess* 7:1–190
26. Newman S, Steed L, Mulligan K (2004) Self-management interventions for chronic illness. *Lancet* 364:1523–1537
27. Norris SL, Engelgau MM, Narayan KM (2001) Effectiveness of self-management training in type 2 diabetes: a systematic review of randomized controlled trials. *Diabetes Care* 24:561–587
28. Mühlhauser I, Sawicki PT, Didurjeit U, Jörgens V, Trampisch HJ, Berger M (1993) Evaluation of a structured treatment and teaching programme on hypertension in general practice. *Clin Exp Hypertens* 15:125–142
29. Mühlhauser I, Sawicki P, Didurjeit U, Jörgens V, Berger M (1988) Uncontrolled hypertension in type 1 diabetes: assessment of patients' desires about treatment and improvement of blood pressure control by a structured treatment and teaching programme. *Diabet Med* 5:693–698
30. Sawicki P, Mühlhauser I, Didurjeit U et al (1993) Optimising structures of antihypertensive therapy—long-term results of a randomised-controlled trial in medical practices [article in German]. *Dtsch Ärztebl* 90:1736–1741
31. Sawicki PT, Mühlhauser I, Didurjeit U, Baumgartner A, Bender R, Berger M (1995) Intensified antihypertensive therapy is associated with improved survival in type 1 diabetic patients with nephropathy. *J Hypertens* 13:933–938
32. Kronsbein P, Mühlhauser I, Jörgens V, Scholz V, Venhaus A, Berger M (1988) Evaluation of a structured treatment and teaching programme on non-insulin-dependent diabetes. *Lancet* 2:1407–1411
33. Grüsser M, Hartmann P, Schlottmann N, Lohmann FW, Sawicki PT, Jörgens V (1997) Structured patient education for outpatients with hypertension in general practice: a model project in Germany. *J Hum Hypertens* 11:501–506
34. Domenech MI, Assad D, Mazzei ME, Kronsbein P, Gagliardino JJ (1995) Evaluation of the effectiveness of an ambulatory teaching/treatment programme for non-insulin dependent (type 2) diabetic patients. *Acta Diabetol* 32:143–147
35. Pieber TR, Holler A, Siebenhofer A et al (1995) Evaluation of a structured teaching and treatment programme for type 2 diabetes in general practice in a rural area of Austria. *Diabet Med* 12:349–354
36. Grüsser M, Bott U, Scholz V, Kronsbein P, Jörgens V (1992) The introduction of a structured treatment and teaching programme for type 2 diabetes in general practices [article in German]. *Diab Stoffw* 1:229–234
37. Gagliardino JJ, Etchegoyen G (2001) A model educational program for people with type 2 diabetes: a cooperative Latin American implementation study (PEDNID-LA). *Diabetes Care* 24:1001–1007
38. Mühlhauser I, Bruckner I, Berger M et al (1987) Evaluation of an intensified insulin treatment and teaching programme as routine management of type 1 (insulin-dependent) diabetes. The Bucharest-Düsseldorf Study. *Diabetologia* 30:681–690
39. Starostina EG, Antsiferov M, Galstyan GR et al (1994) Effectiveness and cost-benefit analysis of intensive treatment and teaching programmes for type 1 (insulin-dependent) diabetes mellitus in Moscow: blood glucose versus urine glucose self-monitoring. *Diabetologia* 37:170–176
40. Jörgens V, Grüsser M, Bott U, Mühlhauser I, Berger M (1993) Effective and safe translation of intensified insulin therapy to general internal medicine departments. *Diabetologia* 36:99–105
41. Bott U, Jörgens V, Grüsser M, Bender R, Mühlhauser I, Berger M (1994) Predictors of glycaemic control in type 1 diabetic patients after participation in an intensified treatment and teaching programme. *Diabet Med* 11:362–371
42. Scholz V, Jörgens V, Berger M et al (1984) Evaluation of a post-graduate course for diabetes educators. *Diabetes Educ* 10:80–84
43. Pieber TR, Brunner GA, Schnedl WJ, Schattenberg S, Kaufmann P, Kreis GJ (1995) Evaluation of a structured outpatient group education program for intensive insulin therapy. *Diabetes Care* 18:625–630
44. Trocha AK, Schmidtke C, Didurjeit U et al (1999) Effects of intensified antihypertensive treatment in diabetic nephropathy: mortality and morbidity results of a prospective controlled 10-year study. *J Hypertens* 17:1497–1503
45. Mühlhauser I, Overmann H, Bender R, Jörgens V, Berger M (2000) Predictors of mortality and end-stage diabetic complications in patients with type 1 diabetes mellitus on intensified insulin therapy. *Diabet Med* 17:727–734
46. DAFNE Study Group (2002) Training in flexible, intensive insulin management to enable dietary freedom in people with type 1 diabetes: dose adjustment for normal eating (DAFNE) randomised controlled trial. *BMJ* 325:746–749
47. Berger M, Mühlhauser I (1999) Diabetes care and patient-oriented outcomes. *JAMA* 281:1676–1678
48. Jörgens V, Grüsser M, Bott U, Mühlhauser I, Berger M (1993) Effective and safe translation of intensified insulin therapy to general internal medicine departments. *Diabetologia* 36:99–105
49. Berger M, Jörgens V, Flatten G (1996) Health care for persons with non-insulin-dependent diabetes mellitus. The German experience. *Ann Intern Med* 124:153–155
50. Grüsser M, Bott U, Ellermann P, Kronsbein P, Jörgens V (1993) Evaluation of a structured treatment and teaching program for non-insulin-treated Type 2 diabetic outpatients in Germany after the nationwide introduction of reimbursement policy for physicians. *Diabetes Care* 16:1268–1275
51. Müller UA, Femerling M, Reinauer KM et al (1999) Intensified treatment and education of Type 1 diabetes as clinical routine. A nationwide quality-circle experience in Germany. ASD (the Working Group on Structured Diabetes Therapy of the German Diabetes Association). *Diabetes Care* 22(Suppl 2):B29–B34
52. Müller UA, Köhler S, Femerling M (2000) HbA1c and severe hypoglycaemia after intensified treatment and education of type 1 diabetes as clinical routine. Results of a nationwide quality-circle in Germany (ASD) 1992–1999 [article in German]. *Diab Stoffw* 9:67–81
53. Bott S, Bott U, Berger M, Mühlhauser I (1997) Intensified insulin therapy and the risk of severe hypoglycaemia. *Diabetologia* 40:926–932
54. Bott U, Mühlhauser I, Overmann H, Berger M (1998) Validation of a diabetes-specific quality-of-life scale for patients with type 1 diabetes. *Diabetes Care* 21:757–769
55. Heinemann L, Overmann H, Mühlhauser I (1997) Quality of blood glucose self monitoring of patients with type 1 diabetes mellitus 6 years after participation in a structured treatment and teaching program [article in German]. *Diab Stoffw* 6:59–63
56. Mühlhauser I, Sawicki PT, Blank M, Overmann H, Bender R, Berger M (2000) Prognosis of persons with Type 1 diabetes on intensified insulin therapy in relation to nephropathy. *J Intern Med* 248:333–341
57. Mühlhauser I, Overmann H, Bender R et al (1998) Social status and the quality of care for adult people with type 1 (insulin-dependent) diabetes mellitus—a population-based study. *Diabetologia* 41:1139–1150
58. Mühlhauser I, Overmann H, Bender R, Bott U, Berger M (1998) Risk factors of severe hypoglycaemia in adult patients with type 1 diabetes—a prospective population based study. *Diabetologia* 41:1274–1282
59. Tankova T, Dakovska G, Koev D (2001) Education of diabetic patients—a one year experience. *Patient Educ Couns* 43:139–145
60. Atkins D, Best D, Briss PA et al (2004) Grading quality of evidence and strength of recommendations. *BMJ* 328:1490–1494