



University of Southern Denmark

A Systematic Review of Instruments Measuring the Division of Care Responsibilities between Children with Type 1 Diabetes and their Parents

Aalders, Jori; Nefs, Giesje; Hartman, Esther; A Nguyen, Linh; Winterdijk, Per; van Mil, Edgar; Aanstoot, Henk-Jan; Pouwer, Frans

Published in:
Current Diabetes Reviews

DOI:
10.2174/1573399818666220510172511

Publication date:
2023

Document version:
Accepted manuscript

Citation for pulished version (APA):

Aalders, J., Nefs, G., Hartman, E., A Nguyen, L., Winterdijk, P., van Mil, E., Aanstoot, H-J., & Pouwer, F. (2023). A Systematic Review of Instruments Measuring the Division of Care Responsibilities between Children with Type 1 Diabetes and their Parents. *Current Diabetes Reviews*, 19(5), 59-93. [e100522204559]. <https://doi.org/10.2174/1573399818666220510172511>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk

Title

A Systematic Review of Instruments Measuring the Division of Care Responsibilities between Children with Type 1 Diabetes and their Parents

Running title

Measuring the division of diabetes care responsibilities

Author information

Jori Aalders^{1,2,3}, Giesje Nefs^{1,4,5}, Esther Hartman¹, Linh A. Nguyen¹, Per Winterdijk⁵, Edgar van Mil⁶, Henk-Jan Aanstoot⁵, Frans Pouwer^{2,3,7}

¹ Center of Research on Psychological and Somatic disorders [CoRPS], Department of Medical and Clinical Psychology, Tilburg University, Tilburg, the Netherlands

² Department of Psychology, University of Southern Denmark, Odense, Denmark

³ Steno Diabetes Center Odense, Odense, Denmark

⁴ Radboud university medical center, Radboud Institute for Health Sciences, Department of Medical Psychology, Nijmegen, the Netherlands

⁵ Diabeter, Center for pediatric and adolescent diabetes care and research, Rotterdam, the Netherlands

⁶ Kidz&Ko, Jeroen Bosch Hospital, 's-Hertogenbosch, the Netherlands

⁷ School of Psychology, Deakin University, Geelong, Australia

Corresponding author

Dr Jori Aalders

Address: Klørvænget 10/13, 5000 Odense C

Phone: +45 65504376

E-mail: jaalders@sdu.dk

ORCID identifiers

Dr Jori Aalders: 0000-0002-7207-8304

Dr Giesje Nefs: 0000-0002-8772-740X

Dr Esther Hartman: 0000-0003-3635-4487

Dr Linh A. Nguyen: 0000-0002-9949-5546

Per Winterdijk, MD: 0000-0003-3352-0523

Prof. Edgar van Mil, MD: 0000-0002-0888-0730

Dr Henk-Jan Aanstoot, MD: 0000-0002-5534-1633

Prof. Frans Pouwer: 0000-0002-8172-9818

ABSTRACT

Background

The division of care responsibilities between parents and children with type 1 diabetes, and an optimal transfer of responsibilities from parent to child over time are assumed to be key for optimal diabetes outcomes during childhood and adolescence. However, an overview of instruments assessing this division as well as their psychometric qualities is currently lacking.

Objective

To 1) identify all existing instruments, 2) evaluate their psychometric properties, and 3) provide an overview of scoring methods.

Methods

Pubmed and PsycINFO were searched using an a priori defined search string. Peer-reviewed studies in English using an instrument assessing the division of diabetes care responsibilities between children (6-18 years) and parents were included. In total, 84 of 725 articles qualified, covering 62 unique samples.

Results

Thirteen questionnaires were identified. The Diabetes Family Responsibility Questionnaire (DFRQ) was most frequently used across studies. Instructions, content and number of tasks, response options, and scoring methods varied across questionnaires. Recent studies often adapted questionnaires, contributing to the heterogeneity across measures. Overall, reporting and quality of psychometric properties was suboptimal.

Conclusion

The division of diabetes care responsibilities can be operationalized with various instruments, each having its strengths and weaknesses but all with limited psychometric support. To measure the division of diabetes care responsibilities more adequately, an updated version of the popular DFRQ or a new scale needs to be developed and evaluated.

KEYWORDS: Type 1 Diabetes Mellitus, Parents, Child, Self-care, Transfer, Psychometrics

1. INTRODUCTION

Despite ongoing technological improvements in glucose measurement and insulin administration, managing pediatric type 1 diabetes is still a complex task for families, requiring intensive 24/7 self-care. Young children with type 1 diabetes are not yet capable of fully understanding and performing all self-care activities [1], and in adolescence motivational problems are common [2]. Therefore, parents often bear responsibility for diabetes care [3]. Similar to facilitating general independence, parents need to guide their children to independent diabetes care in adulthood by transferring these care responsibilities to their child over time.

How diabetes tasks are divided between parents and children is an important issue to monitor and discuss in clinical practice as the timing of transfer moments can impact self-care activities [4, 5], glycemic outcomes [5], the child's global self-worth and self-efficacy [5], diabetes-specific quality of life [6], and the success of transition into adult health care [7]. Moreover, children and parents can have conflicting opinions regarding who is responsible for diabetes care tasks (e.g., the parent reports that the child is responsible while the child reports that the parent is responsible) [8, 9]. Disagreement about who is responsible for diabetes care tasks within parent-child dyads has been associated with suboptimal glycemic outcomes [8-13], higher diabetes-specific distress [12] and more diabetes related conflicts [14].

For clinicians and researchers, different instruments have been developed to measure the division of diabetes care responsibilities between parents and their child. Furthermore, across studies diverse scoring methods have been reported for processing raw instrument data into final scores.

In order to better understand how diabetes care responsibilities are divided within families and to examine changes in this division over time, it is important to select the most suitable instrument and scoring method after thorough consideration of the psychometric properties of all available instruments. Currently, there is no systematic review available of existing instruments assessing the division of diabetes care responsibilities nor a structured evaluation of their psychometric qualities.

To guide healthcare providers and researchers in selecting an instrument and scoring method that best fit their research or clinical objectives and to get a better understanding of how this construct is conceptualized across studies, this systematic review aims to: a) identify instruments that have been used in prior studies to measure the division of diabetes care responsibilities between parents and children, b) evaluate the reported psychometric properties of these instruments, and c) provide an overview of scoring methods used to process raw instrument data into final scores.

2. METHODS

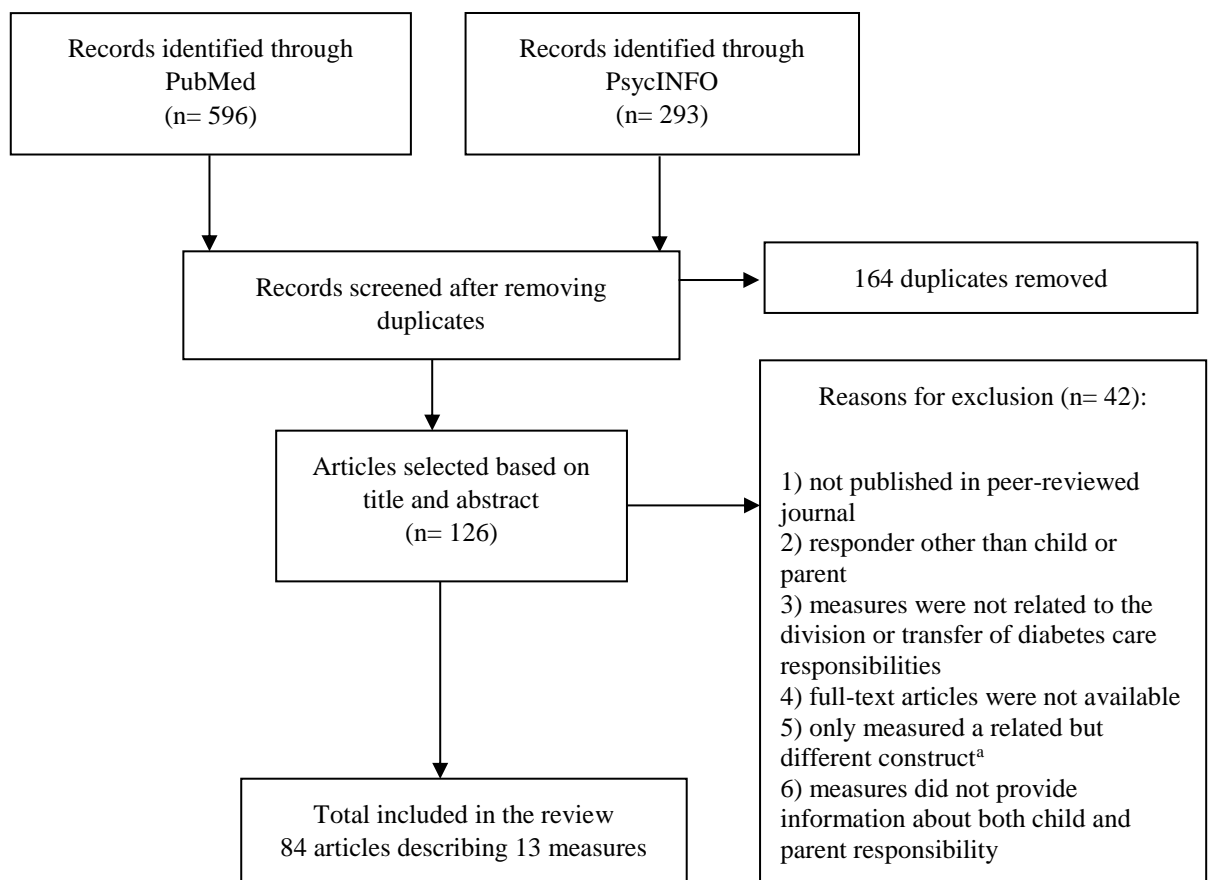
2.1. Data sources and searches

On June 1st 2018 a systematic search was performed in PubMed and PsycINFO. Given the rapid development of new diabetes self-care devices and human insulins since the 1980's, the search was limited to studies published from January 1st 1990 onward. Titles, abstracts and keywords were scanned using a combination of search terms (Table 1). The search strategy resulted in 596 hits in PubMed and 293 in PsycINFO. After removal of duplicate studies, 725 unique articles remained (Figure 1).

Table 1. Search terms.

PubMed
<p>1#: (“diabetes mellitus type 1” [tiab] OR “diabetes mellitus type I” [tiab] OR “diabetes type 1” [tiab] OR “diabetes type I” [tiab] OR “type 1 diabetes” [tiab] OR “type I diabetes” [tiab] OR “insulin dependent diabetes” [tiab] OR “insulin-dependent diabetes” [tiab] OR IDDM [tiab] OR “juvenile diabetes” [tiab] OR “juvenile-onset diabetes” [tiab] OR “juvenile onset diabetes” [tiab] OR “diabetes mellitus, type 1”[MESH])</p> <p>2#: (juvenile* [tiab] OR adolescen* [tiab] OR pubert* [tiab] OR teen* [tiab] OR youth* [tiab] OR young* [tiab] OR child* [tiab] OR pediatric [tiab] OR paediatric [tiab] OR “child” [MESH] OR “adolescent” [MESH])</p> <p>3#: (parent* [tiab] OR caregiv* [tiab] OR father* [tiab] OR mother* [tiab] OR maternal [tiab] OR paternal [tiab] OR “parents” [MESH])</p> <p>4#: (responsib* [tiab] OR autonom* [tiab] OR involv* [tiab] OR transfer* [tiab] OR "transference (psychology)"[MeSH Terms])</p> <p>1# AND 2# AND 3# AND 4# Filter results: 01-01-1990 to 01-06-2018</p>
PsycINFO
<p>(“diabetes mellitus type 1” OR “diabetes mellitus type I” OR “diabetes type 1” OR “diabetes type I” OR “type 1 diabetes” OR “type I diabetes” OR “insulin dependent diabetes” OR “insulin-dependent diabetes” OR IDDM OR “juvenile diabetes” OR “juvenile-onset diabetes” OR “juvenile onset diabetes”) AND (juvenile* OR adolescen* OR pubert* OR teen* OR youth* OR young* OR child* OR pediatric OR paediatric) AND (parent* OR caregiv* OR father* OR mother* OR maternal OR paternal) AND (responsib* OR autonom* OR involv* OR transfer*)</p> <p>Filter results: 01-01-1990 to 01-06-2018</p>

Figure 1. Flow chart of the search strategy.



^a These studies only measured a construct that was related to the division of diabetes care responsibilities (i.e., parental monitoring, parental supervision, appraisal of parental involvement, parental autonomy supporting behavior, decision making process between dyads about diabetes care, collaborative involvement, parental knowledge and child disclosure with regard to diabetes care, parental expectations of the division, transition readiness, and child mastery of diabetes care tasks).

2.2. Study selection

Two reviewers (JA and GN) independently screened titles and abstracts of the unique hits based on the following inclusion criteria: 1) written in English, 2) study population consisted of children with type 1 diabetes (6-18 years) and/or their parents, 3) the division or transfer of diabetes of diabetes care responsibilities within families was assessed (e.g., studies focusing on the transition from a pediatric to an adult diabetes clinic were not included), 4) studies were published in a peer-reviewed journal, 5) studies had a quantitative design (e.g., studies merely presenting qualitative and review findings were excluded), and 6) assessments were completed by parents and/or children (i.e., health care professionals reports were excluded). A 98% agreement-rate was obtained. Because of initial non-agreement about inclusion, 16 (2%) full-text articles were discussed until consensus was reached [15-30].

During full-text screening, studies that merely included instruments assessing a construct related to the division of diabetes care responsibilities were excluded (see Figure 1 for a detailed description). Additionally, studies were excluded when items regarding the division of diabetes care responsibilities were combined with items measuring related but not identical constructs to obtain (sub)scale scores. Furthermore, studies were excluded if only the frequency of parental instrumental assistance in diabetes care or the division of diabetes care tasks between parents was examined, as these measures do not fully capture how diabetes care tasks are divided between children and parents.

If selected studies were (partly) based on the same sample, data concerning reliability was included from the most recently published study as this information was considered as most up to date. It was checked whether older studies contained additional information warranting inclusion.

2.3. Data extraction and quality assessment

The first author (JA) extracted relevant data from included studies. For each individual study, data was collected on: study and sample characteristics (i.e., country of data collection, design, responder(s), number of participants and information about the child's gender, age, ethnicity, household composition, HbA_{1c}, pump use, and diabetes duration) and the included instrument measuring the division of treatment responsibilities between parents and children (i.e., number of items, response options, scoring procedure, potential range, final scores, psychometric information). If these characteristics were reported separately for two or more subgroups, total sample scores were computed if possible [31]. Other authors (GN, EE, LN, FP) independently checked extractions. Study quality was evaluated in terms of data availability for extraction of study, sample and instrument characteristics.

2.4. Data synthesis and analysis

The evaluation of psychometric properties was based on the COSMIN guideline [32-34]. In line with this guideline, the evaluation of content validity (i.e., relevance, comprehensiveness, comprehensibility) was based on

the findings and the methodological quality of (a) the original study (i.e., describes the development of the scale or the first article that introduces the instrument) and (b) the content validity study (i.e., examines the relevance, comprehensiveness and comprehensibility of an existing scale). Moreover, the reviewers' own observations are included in this evaluation (for more details see appendix A and the COSMIN guideline). As instructed, structural validity (factor structure), internal consistency (Cronbach's alpha), inter-rater reliability and construct validity (i.e., convergent, concurrent, discriminative or known-group validity) of the identified questionnaires were evaluated by (a) rating the measurement properties, (b) rating the design and methodological quality at the study level and, if applicable, (c) assessing the quality of evidence (level of inconsistency, impression and indirectness across studies) (see Appendix B and C for criteria). As criteria to evaluate the methodological quality of studies using exploratory factor analysis to examine structural validity are limitedly defined within the COSMIN guideline, the checklist for exploratory factor analysis as suggested by Howard, 2016 was included to evaluate methodological quality of these studies [35]. As changes in the division of responsibilities over time are considered as normal [36, 37], test-retest reliability was only included if follow-up time was limited to one month. Hypothesis testing for construct validity was examined if it was explicitly described that comparisons with other instruments or between subgroups were conducted in the context of construct validity research. Criterion validity was not included in the analysis, given there is no gold-standard for measuring the division of diabetes care responsibilities. Feasibility was assessed by examining completion time, possible costs and the availability in different languages.

Based on both standard checklists [32-35], a narrative evaluation of psychometric properties was performed to facilitate recommendations about which existing instrument to use when taking psychometric properties into account. In addition, the COSMIN guideline was used to generate an overall quantitative evaluation per psychometric property for each instrument and to evaluate the overall quality of evidence that was included in this evaluation.

3. RESULTS

3.1. Study and sample characteristics of the included studies

In total, 84 articles [4-6, 8-14, 36-109], were identified that used an instrument (all questionnaires) to assess (some aspect of) the division of diabetes care responsibilities in 62 unique samples. Table 2 provides a summary of the sample characteristics of all studies per questionnaire. Studies that were (partly) based on same samples and reported about the same questionnaire are labelled as such. Most data were (at least partly) collected in the USA (84%, 52 out of 62 unique samples) [4-6, 8-11, 13, 14, 36-42, 44-49, 51, 53-70, 72-80, 83, 85-93, 95-104, 106-109]. Among the 84 individual studies, the minority reported about assessments across multiple time points as part of an intervention study (n= 10)[54, 63, 66, 77, 78, 81-83, 94, 105] or had a longitudinal observational design (n= 11) [4-6, 62, 68, 69, 91, 92, 97, 103, 107]. Most studies included both parent- and child-reported responsibility (n= 54) [4, 5, 8-14, 36-39, 41-44, 46-48, 51, 53, 55-57, 60-67, 70, 72, 73, 75, 76, 79-81, 83, 85, 87-89, 91, 96, 98, 101-104, 106], while 26 studies only reported about the assessment of one family member (n parents= 18 [49, 50, 54, 68, 71, 74, 78, 90, 92-95, 99, 100, 105, 107-109], n children= 8 [6, 40, 58, 59, 69, 77, 86, 97]); for three studies it was unclear who reported data [52, 82, 84] and for one study it was unclear if, next to parental reports, children also completed the responsibility measure [45]. Two articles, based on the same sample, included two instruments

simultaneously [92, 107]; in this sample separate assessment tools were used for children using an insulin pen and pump [92, 107].

3.2. Retrieved instruments and frequency of use

In total, 13 different questionnaires were identified measuring the division of diabetes care responsibilities between parents and children [8, 40, 50, 52, 58, 78, 84, 91, 100, 107, 110-112]. When considering studies (partly) based on the same sample as one study, the Diabetes Family Responsibility Questionnaire (DFRQ) was by far the most frequently used (n= 42 studies) [5, 8-12, 36, 37, 42, 45-49, 51, 53, 60-67, 70, 72-77, 79-83, 85-87, 93, 95, 96, 98, 99, 101, 105, 106, 108, 109], followed by the Diabetes Responsibility and Conflict Scale (DRCS) (n= 7) [4, 6, 38, 39, 41, 43, 54-57, 71, 88, 89, 94, 97, 102-104]. The following ten instruments were infrequently used (1-2 unique samples): Continuous Subcutaneous Insulin Infusion (CSII)-Use survey [92, 100, 107], Independent functioning in daily and nondaily diabetes management checklist [58, 59], Deciding About Diabetes Treatment Scale (DADT) [13, 14], Diabetes Regimen Responsibility Scale (DRRS) [44, 90], adapted child version of the Diabetes Independence Survey (DIS) [40], adapted parent version of the Diabetes Independence Survey (DIS) (pen users) [92, 107], Children's Diabetes Inventory (CDI) [78], and nameless instruments described by Friedemann-Sanchez et al. [50], Nakamura et al. [84], Gowers et al. [52], and Robinson et al. [91].

3.3. General characteristics of retrieved instruments

Table 3 presents the general characteristics of these instruments as described in the original study and provides information about how these instruments have been used in subsequent studies. As the search was limited to peer-reviewed articles that were published from 1990 onwards, three older original studies had to be assessed separately [110-112]. As these three development studies provided limited information (i.e., conference abstract, thesis and article that were not accessible) [110-112], information was complemented by consistently reported information from later studies using these instruments.

3.4. Instrument type

Among the instruments, only questionnaires were identified. One questionnaire consisted of a combination of items from two other included questionnaires [91]. Two scales were based on the DIS [113, 114], that was originally developed to assess the child's mastery level of diabetes care skills; separate adaptations in the instructions and answer options were made for parents and children in order to measure the division of diabetes care tasks [40, 107]. For four instruments, the assessment of the division of diabetes care responsibilities was a separate subscale of an instrument that also assessed other constructs for the included tasks (e.g., child's mastery level of diabetes related skills [100], decision making about who performs tasks [58], behavioral frequency [78], conflict [110]).

3.5. Instrument target population and origin

Five instruments were developed to assess both parent and child perceptions of the division of diabetes care tasks [8, 91, 110-112]. Eight instruments were developed specifically for parents [50, 52, 78, 100, 107] or children [40, 58, 84]. The majority of instruments was developed in English (85%) [8, 40, 52, 58, 78, 91, 100, 107, 110-112] and originated from the USA (77%) [8, 40, 58, 78, 91, 100, 107, 110-112]. Six scales were developed before 2000 [8, 52, 78, 110-112].

3.6. Instrument general structure, instructions and recall period

All instruments requested responders to indicate per individual task (items/questions) how tasks/responsibilities were divided between parents and children (response options of questionnaires). Instructions and response wording differed across instruments; five instruments focused on behaviors related to the actual performance of diabetes care tasks (who performed, decided about or oversaw tasks) [52, 58, 110, 112] whereas six instruments referred to the division of responsibilities (who takes or initiates responsibility, is responsible for monitoring/ensuring a task was completed) [8, 40, 78, 91, 100, 107]. For two instruments instructions were not available [84, 111]. Only for one instrument recall periods were specified; responders were requested to base their answers on the previous day for daily tasks and the previous three months for non-daily tasks [58].

3.7. Instrument content (type of tasks; pen vs. pump tasks; level of tasks) and number of included tasks

Considerable differences were observed regarding instrument content and number of included tasks. One instrument appeared to be limited to administering insulin [52], whereas the other original instruments included a broader range of diabetes care tasks, such as exercising, monitoring general health, managing supplies, contact with diabetes care team, discussing diabetes with family and friends [8, 40, 50, 58, 78, 84, 91, 100, 107, 110-112]. Furthermore, almost all original instruments did not describe tasks specific for children using a pump [8, 40, 50, 52, 58, 78, 84, 91, 107, 110-112]. In contrast, the CSII-Use survey specifically focused on self-care tasks related to pump therapy (e.g., programming basal rate) [100]. Additionally, task levels varied across instruments; based on descriptions, example items and response options, one instrument merely focused on decision making [112], two instruments focused on performing tasks [52, 91], while four instruments mixed items about performing tasks, making decisions and thinking about tasks [8, 50, 78, 110]. For six instruments, the level of included tasks was unclear [40, 58, 84, 100, 107, 111]. The number of included tasks in the original instrument varied from 15 to 42 (n= 10) [8, 40, 58, 78, 91, 100, 107, 110-112]. For one instrument it was unclear whether the assessment was limited to one item regarding insulin injections or whether multiple items were included next to this item [52].

3.8. Instrument response scales

To indicate how tasks or responsibilities were divided between parents and children, items of questionnaires were, according to the original studies, rated on a dichotomous scale (“yes”) [84], a 3-point scale (parent is responsible, child is responsible, shared responsibility) [8, 40, 58, 100, 107], and a 5-point scale (ranging from full child responsibility to full parent responsibility) [78, 91]. For one scale response options differed as responders were requested to indicate who in the family was primary responsible (answer options: child, mother, father, or other) [50]. While a few original studies did not report response options [52, 110-112], studies that were published later consistently defined response options on a 5-point scale for the DRCS [110] (one end= child or parent, middle= shared) [4, 6, 38, 39, 41, 43, 55-57, 71, 88, 89, 97, 102-104] and a 4-point scale for the DADT [112] (no decision-parent decides, discussion-parent decides, discussion-child decides, no discussion-child decides) [13, 14]. For one measure, only two response options were stated in a table in the result section of the original paper (child gives injections alone, parent and child give injections together), however, based on this information it is unclear if more answer options were included [52]. For one measure [111], inconsistent response options were defined in later studies using this scale [44, 90].

Among the reported response options, seven instruments included a separate option to indicate that responsibility was shared between parents and children [8, 40, 52, 58, 100, 107, 110]. Furthermore, two instruments included answer options to indicate that the task is not relevant within the family (“no one”) or within the division between

parents and children (“other”) [50, 58]. Five original studies defined values for response options [8, 40, 58, 78, 107]; four instruments assigned consecutive values (starting at “1”) to each answer category with higher values indicating more child responsibility [8, 58, 78, 107] or more parent responsibility [40]. One instrument assessing the child’s level of independent functioning, assigned a value of “2” when tasks were reported as the child’s responsibility whereas other answer options (i.e., parent, shared or no one’s responsibility) were coded with a value of “1” [58].

3.9. Use of instruments across included studies – deviations from original versions

Differences were observed regarding how instruments were described in original studies and how these instruments were used in later studies. For example, in some later studies various subsets and adapted and/or updated items were generated that differed from the original tool [4, 6, 10, 38, 39, 41, 43, 45, 47, 68, 71, 81, 83, 88, 89, 92, 94, 95, 97, 102-104, 107] or response options were added (i.e., “not applicable”, “someone else” or “no one in the home”) [5, 47, 63, 83]. Contrary to original instructions of the DFRQ [8], in one study responders were requested to report per item which family member had most responsibility instead of who takes or initiates responsibility [47]. Furthermore, in another study instructions were adapted to be inclusive for both pump and pen users [55]. Finally, in 12 studies the value labels that were described in the original study were modified (i.e., values of answer categories were reversely assigned, scoring starts at 0 instead of 1) [9, 12, 37, 46, 59, 62, 67, 75, 79, 82, 96, 98]. This comparison was limited to 23 studies as 27 studies did not report how response options were coded [5, 10, 26, 36, 37, 42, 47, 48, 51, 53, 60, 63-66, 70, 72, 74, 77, 80, 81, 83, 85, 86, 99, 101, 106, 108, 109, 115, 116]. Instruments whose value labels were not defined in the original study also showed inconsistencies regarding the coding of response options across studies [4, 6, 38, 39, 41, 43, 55-57, 71, 88, 89, 97, 102-104]; in addition to variations in starting points of values (0 or 1) and whether values were assigned to indicate increasing child or parent responsibility, one study coded response options in a way that higher item scores indicated more shared responsibility [55].

3.10. Scoring methods

Diverse scoring methods were reported to process raw instrument data into final scores. Instead of a responsibility score per individual responder, a few studies generated a family responsibility score by averaging parent- and child-reports [60, 61, 63, 91]. Furthermore, some studies reported item scores [45, 52, 73] or subscales scores [8, 44, 45, 50, 60, 75, 78, 90, 98, 100]. Finally, different calculations were used to generate (sub)scale scores; most studies generated sum [10, 14, 37, 40, 44-46, 48, 49, 54-60, 62-65, 67, 71, 74, 76, 79, 81, 82, 84-87, 90, 94-96, 98, 99, 101, 106, 108] or mean scores [4-6, 11, 38, 41, 43, 66, 75, 78, 88, 89, 91, 92, 102-105, 107] of items, with higher scores reflecting more parent [4-6, 10, 37, 38, 40, 41, 43, 46, 56, 57, 62-64, 67, 71, 75, 79, 81, 82, 84, 85, 88, 89, 91, 96, 98, 101-104, 106, 108], child [11, 45, 48, 49, 57-60, 65-67, 74, 78, 86, 87, 90, 92, 95, 99, 105, 107] or shared responsibility [55]. Additionally, some studies contrasted the level of solely parent responsibility, shared responsibility or solely child responsibility by adding the number of tasks [47, 50] or calculating the percentage of tasks per answer category [5, 49, 50, 52, 61, 76, 100]. For 12 studies, it was unclear how responsibility summary scores were obtained [8, 36, 39, 42, 68-70, 77, 80, 93, 97, 109].

As it is complex to evaluate responsibility scores given the limited availability of guidelines indicating the “right” division of diabetes care responsibilities, scores were in some studies evaluated by subsequently assigning families to categories according to sample distributions (i.e., tertiles) [37, 74], mean or sum (sub)scale scores [107], or to

the response option that was most frequently filled in [45]. Additionally, in one study a ratio score between a responsibility measure and another measure was used to assess the extent of “appropriate” child responsibility [109].

In total, 12 studies generated scores for the extent of agreement and/or disagreement within a parent-child dyad about the division of responsibility [9-14, 51, 53, 72, 76, 82, 83]. As this approach is focused on whether it is likely that diabetes care responsibilities are taken or generate conflict within dyads rather than on the division of responsibilities, these scoring methods are not discussed in further detail.

3.11. Reported psychometric properties of the instruments across the studies that used the instrument

Limited information was available regarding the psychometric properties of the included instruments (see Table 4 for a narrative evaluation and Table 5 for the overall quantitative rating). As no relevant test-retest (<one month) and inter-rater reliabilities were reported, these aspects are not discussed.

3.11.1 Content validity

Regarding content validity, for ten instruments information regarding the development of the scale was available [8, 40, 50, 58, 78, 84, 91, 100, 107, 110]. Information regarding the relevance of included items for the construct of interest was for one instrument limited to a description of the target population [84] and for two instruments limited to a description of the target population and the construct of interest [50, 110]. Although nine instruments provided a description of the construct of interest, it often remained unclear how the construct of interest (i.e., responsibility [8, 40, 78, 91, 107, 110], autonomy [100]) was defined. Furthermore, none of the original studies described an underlying theoretical model or how the concept is theoretically related to other concepts. For the instruments that were (partly) based on concept elicitation [8, 100], previous developed scales [40, 58, 91, 107] or that were reviewed by professionals [58, 78], no evidence was provided that indicated to what extent included items were relevant to the construct of interest.

Accordingly, information regarding the relevance of items for the target population was limited; for most questionnaires the target population was not explicitly stated, but rather deduced from the population in the original study [40, 50, 58, 78, 84, 91, 107, 110]. Only for the DFRQ it was described that the target population was represented in the elicitation of relevant items [8]. However, only mothers, not fathers, were represented in the parent sample and methodological information regarding concept elicitation was limited to the description of the type of data collection [8]. Merely a few studies provided information to evaluate the context of interest [8, 78, 100] and the justification of included response options was limited to references to other scales that did not justify why particular response options were included [40, 107].

In none of the original studies, information regarding comprehensiveness of the instrument by the population of interest was provided. Based on the description of items and the years in which these questionnaires were developed, it is likely that most questionnaires consist of outdated tasks and do not include tasks that are related to new technologies [8, 40, 50, 52, 58, 78, 84, 91, 110-112]. Comprehensibility was only examined for the CSII-Use survey [100]. However, results and methodological information of this comprehensibility study remained unclear.

Reviewer ratings for comprehensibility were often based on descriptions of scales and example items as full questionnaires were not available. Scales using the terms “responsibility” [8, 40, 78, 91, 100, 107] and “shared

responsibility” [8, 40, 58, 100, 107, 110] might provide insufficient clarification about the interpretation of these concepts for those completing them; responsibility can refer to both behavioral and emotional aspects and shared responsibility can be interpreted in terms of frequency as well as in terms of shared performance/decision making. With regard to response options, only two instruments included an option to indicate that the listed tasks were not applicable for the division of diabetes care responsibilities/tasks between parents and children [50, 58].

Overall, aspects of content validity were rated as insufficient or indeterminate. The quality of evidence was rated as (very) low.

3.11.2 Internal structure and reliability

For six questionnaires, subscales were defined in original studies [8, 50, 58, 84, 100, 111] and later studies [60, 98]. Only for the DFRQ [8], factor analysis was used to develop the scale [8] and to examine the internal structure of the final instrument [98]. In the original study, no a priori hypotheses were formulated regarding the structure of the scale [8]. For maternal reports in this study, explanatory factor analysis identified three subscales (methods unclear), named: “general health maintenance”, “regimen tasks” and “social presentation of diabetes” [8]. For child reports, no such structure emerged [8]. In a later study, explanatory factor analysis was used to explore whether the DFRQ could be divided into direct and indirect tasks [98]. Authors concluded that for both parent- and child-reports, explanatory factor analysis supported this hypothesis [98]. However, according to a standard checklist for explanatory factor analysis, these conclusions are not fully supported and both studies had important methodological shortcomings (see Table 4). For the CSII-Use survey assessing both the division of responsibilities and the mastery for a set of CSII related tasks, it was unclear if the factor analysis used to design this questionnaire was based on either responsibility or mastery scores [100]. For the only study examining the structural validity of the final instrument (i.e., DFRQ) [98], structural validity was rated as indeterminate as the required criteria were not reported and it was unclear whether the instrument was based on a reflective or formative model, which makes it unclear if structural validity is of concern. The quality of evidence was rated as moderate.

Results regarding internal consistency were mixed (children= 31 studies representing 29 unique samples [4, 8, 10-14, 36, 40, 42-44, 47, 48, 56, 58, 59, 61, 64, 66, 67, 70, 75, 83, 84, 86, 91, 97, 98, 101, 104]; parents= 35 [4, 8, 10-14, 36, 42-44, 47-49, 54, 56, 61, 64, 66, 67, 70, 71, 75, 78, 83, 90-92, 95, 98, 99, 101, 104, 105, 108]). Two questionnaires could not be included in the analysis as no information regarding internal consistency was available [50, 52]. For total scales, low internal consistencies (Cronbach’s alpha <0.70) were only reported for the DFRQ (children= 3 [47, 66, 101] out of 18 [8, 10-12, 36, 42, 47, 48, 61, 64, 66, 67, 70, 75, 83, 86, 98, 101]; parents= 2 [64, 66] out of 22 [8, 10-12, 36, 42, 47-49, 61, 64, 66, 67, 70, 75, 83, 95, 98, 99, 101, 105, 108]). Small sample sizes (n= 30-45) [47, 64, 66], modifications to instructions and response options [47] and paternal reports [64] could have contributed to low internal consistency.

For certain subscales of the DFRQ and the DRRS [8, 111], low internal consistencies were described for both child- and parent-reports [44, 75, 98]. Inconsistencies could be a result of an insufficient factor structure/poor structural validity [44, 75, 98], low number of items in the scale, or could reflect that items are outdated [75, 98].

For all instruments for which internal consistency was examined, internal consistency was rated as indeterminate as the criterium of “at least low evidence for sufficient structural validity” could not be met. The quality of

evidence ranged from moderate to (very) low as no information was available about the structural validity or unidimensionality of the instruments.

3.11.3 Construct validity

Only for the DFRQ and the CDI hypotheses about construct validity were tested [8, 78]. For the DFRQ, mothers who rated their child as more responsible for regimen tasks and across all included tasks also reported independence as a priority for individual family members [8]. However, measurement properties of the comparator instrument were not reported and the analyses were limited to maternal reports. Furthermore, hypotheses regarding associations with (sub)scales and the direction and magnitude of the effect were not predefined. For the CDI, construct validity was examined by known-group validity; parents of younger children (8-10 years) significantly reported less child responsibility than parents of older children (11-12 years) and in correlation analysis increasing age was associated with more overall child responsibility [78]. However, characteristics of the age groups and statistical methods were unclear and hypotheses were not predefined.

Given the lack of predefined hypotheses, construct validity could not be rated. Nevertheless, the quality of the provided evidence would have been (very) low.

3.12. Feasibility

Information regarding feasibility was limited; for none of the included instruments the completion time was stated, however, for most scales the number of items was reported (Table 2) [8, 40, 58, 78, 91, 100, 107, 110-112]. For five scales, both parent and child versions were developed (Table 2) [8, 91, 110-112]. For none of the instruments costs were reported [8, 40, 50, 52, 58, 78, 84, 91, 100, 107, 110-112]. Next to the language in which the questionnaire was developed, a few scales were (partly) translated into different languages in later studies (i.e., DRCS: Chinese [71]; DFRQ: Spanish [61] and presumably in the languages in the 19 countries involved in the study of Cameron et al. 2008 [10]).

Table 2. Summary of sample characteristics per instrument. ^a

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
<i>Diabetes Family Responsibility Questionnaire (DFRQ)</i>											
Agarwal et al. (2016) [36]	USA	Cross-sectional	100 families: 100 children, 100 parents (84% mothers)	43% (43)	13±2.1, range 10-17	37% (36) (97/100)	NR	8.8±1.5, range 6.4-13.6	73±NR, range 46-125	39% (39)	5.9±3.6, range 1.1-14.7
Anderson et al. (1990) [8]	USA	Cross-sectional	121 families: 121 children, 121 parents (100% mothers)	46% (56)	13.3±3.3, range 6-21	12% (14)	87% (105)	11.6±2.7, range 6.3-19.3	103±NR, range 45-187	0% (0)	5.5±3.5, range 1-15
Anderson et al. (2002) [37]	USA	Cross-sectional	104 families: 104 children, 104 parents (77% mothers)	53% (55)	12.1±2.3, range NR, target group 8-17	NR	NR	8.3±1.1, range NR	67±NR, range NR	1% (1)	2.6±1.6, range NR
Anderson et al. (2009) [9] ^A	USA	Cross-sectional	121 families: 121 children and 121 parents ^k (92% mothers, 7% fathers, 1% stepmother, 1% grandmother)	50% (61)	12.1±1.6, range NR, target group 9-14.5	29% (35)	NR	8.4±1.4, range NR, target group <13	68±NR, range NR, target group <119	NR	5.4±3.1, range NR
Caccavale et al. (2015) [42]	USA	Cross-sectional	151 families: 151 children and 151 parents (sex parent NR)	52% (78)	15.6±1.5, range NR, target group 13-18	9% (14)	NR	8.8±1.6, range NR	73±NR, range NR	65% (98)	7.7±3.6, range NR
Cameron et al. (2008) [10]	19 countries ^l	Cross-sectional	2062 families: 2062 children, 1973 parents (77% mothers, 21% fathers, 1% step-parents, 1% guardians)	50% (995) (1973/2062)	14.4±2.05, range NR, target group 11-18 (1973/2062)	NR	79% (1559) (1973/2062)	8.2±1.4, range NR (1973/2062)	66±NR, range NR (1973/2062)	NR	6.1±3.5, range NR (1973/2062)
Cemeroglu et al. (2016) [45]	USA	Cross-sectional	140 families: unclear if children completed responsibility measure, 140 parents ^k (sex parent NR)	NR	10.7±2.0, range 6.2-13.9	NR	NR	8.5±1.1, range 5.8-11.8	69±NR, range 40-105	54% (76)	4.3±2.8, range 0.6-12.2

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
Cortina et al. (2010) [46] ^B	USA	Cross-sectional	150 families: 150 children, 150 parents (87% mothers, 10% fathers, 3% other caregivers)	49% (74)	15.5±1.4, range NR, target group 13-18	13% (20)	75% (113)	8.8±1.9, range NR	73±NR, range NR	63% (95)	6.0±3.9, range NR
Dashiff et al. (2003) [47]	USA	Cross-sectional	31 families: 31 adolescents, 31 mothers, 31 fathers	58% (18)	13.8±1.0, range 12-15	23% (7)	100% (31)	9.4±1.8, range NR	79±NR, range NR	NR	NR
Delamater et al. (2017) [48]	USA	Cross-sectional	51 families: 51 children, unclear if one parent or both parents completed responsibility measure (sex parent NR)	49% (25)	13.5±1.3, range NR, target group 11-16	100% (51)	57% (29)	8.8±1.7, range NR	72.6±NR, range NR		6.0±3.9, range NR
Drotar & Ievers (1994) [49]	USA	Cross-sectional	26 families: 26 parents (100% mothers)	38% (10)	9.6±2.9, range NR, target group 4-14	8% (2)	85% (22)	8.7±1.1, range NR	72±NR, range NR	NR	3.2±0.5, range NR
Geffken et al. (2008) [51]	USA	Cross-sectional	100 families: 100 children, unclear if one parent or both parents completed responsibility measure (sex parent NR)	45% (45)	13.2±2.5, range 7-18	24% (24)	73% (73)	9.0±1.9, range NR	75±NR, range NR	NR	5.6±3.5, range NR
Grabill et al. (2010) [53] ^C	USA	Cross-sectional	224 families: 224 children, 224 parents (sex parent NR)	46% (103)	13.82±2.6, range 8-18 ^m	24% (54)	70% (157)	8.8±1.8, range NR ^m	73±NR, range NR	NR	5.2±3.7, range NR
Helgeson et al. (2008) [5]	USA	Longitudinal	132 families: 132 children, 132 parents (sex parent NR) ^m	47% (62)	12.1±NR, range 10.7-14.2 ^m	7% (9)	NR	8.0±1.3, range NR	64±NR, range NR	26% ^m (34)	4.9±3.0, range 1-13
Holmes et al. (2006) [60]	USA	Cross-sectional	222 families: 222 children, 222 parents (81% mothers)	53% (118)	12.8±1.9, range 9-17	25% (56)	85% (189)	8.3±1.5, range 5.2-15	67±NR, range 33-140	NR	4.2±3.3, range 0.1-13.3
Hsin et al. (2010) [61]	USA	Cross-sectional	111 families: 111 children, 111 parents ^k (83% mothers, 14%	47% (52)	13.3±2.8, range NR, target group 10-17	100% (111)	NR	2.9±2.2, range 4.2-15.4 ^p	NR±NR, range NR ^p	27% (30)	6.2±3.6, range NR

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
			fathers, 3% other guardian)								
Ingerski et al. (2010) [62] ^B	USA	Longitudinal	147 families: 147 children, 147 parents ^k (86% mothers)	48% (71)	15.5±1.4, range NR, target group 13-18 ^m	14% (21)	76% (111)	8.8±1.9, range NR ^m	73±NR, range NR ^m	63.3% (93) ^m	6.0±3.8, range 0.5-16.8
Katz et al. (2014) [63]	USA	Intervention	153 families: 153 children, 153 parents (sex parent NR)	44% (67)	12.9±2.3, range 8.2-16.5, target group 8-16 ^m	9% (14)	NR	8.4±1.4, range NR ^m	68±NR, range NR ^m	23% (35) ^m	6.1±NR, range 0.8-16.5 ^m
Kaugars et al. (2011) [64] ^P	USA	Cross-sectional	69 families: 69 adolescents, 66 mothers, 44 fathers	49% (34)	14.2±1.6, range 12-17	13% (9)	81% (56)	8.2±1.2, range 5.6-12.7	66±NR, range NR	NR	5.5±3.6, range 0.5-14
Kichler et al. (2010) [65] ^D	USA	Cross-sectional	69 families: 69 adolescents, 65 mothers, 45 fathers	51% (35)	14.2±1.6, range 12.0-17.9	13% (9)	81% (56)	8.3±1.3, range 6.0-13.9	67±NR, range 42-128	NR	5.5±3.5, range 0.5-10.6
Kichler et al. (2013) [66]	USA	Intervention	30 families: 30 children, 30 parents ^k (93% mothers)	47% (14)	15.2±1.3, range NR, target group 13-17 ^m	23% (7)	83% (25)	10.0±2.1, range 5.9-14.0 ^m	86±NR, range 41-130 ^m	NR	5.7±3.2, range NR ^m
Kichler et al. (2018) [67]	USA	Cross-sectional	258 families: 258 children, 258 parents (84% mothers)	50% (130)	14.9±1.1, range NR, target group 13-16	78% (200)	NR	9.6±1.2, range NR	81±NR, range NR	70.7 (181)	6.4±3.8, range NR
Lancaster et al. (2015) [11]	USA	Cross-sectional	64 families: 64 children, 64 parents (95% mothers)	50% (32)	13.9±2.8, range 8-18	NR	77% (49)	8.9±1.9, range 5.2-14.5	74±NR, range 33-135	NR	NR
Landers et al. (2016) [70]	USA	Cross-sectional	117 families: 117 children, 117 parents (84% mothers)	44% (51)	12.9±2.5, range 8.1-17.0, target group 8-16	39% (44) (114/117)	NR	8.8±1.5, range 6.1-13.6	73±NR, range 43-125	NR	5.6±3.5, range 1.1-14.7
Law et al. (2013) [12]	UK	Cross-sectional	203 families: 203 children, 203 parents ^k (82% mothers - 129/204)	46% (93)	14.5±1.8, range 11.6-18.8	NR	NR	9.5±1.6, range NR	80±NR, range NR	6% (12)	5.8±3.7, range NR

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
Lehmkuhl et al. (2009) [72] ^c	USA	Cross-sectional	165 families: 165 children, 165 parents (76% mothers, 15% fathers, 4% grandmothers, 4% other caregivers)	41% (67)	13.6±3.0, range 6-18, target range 8-18	28% (47)	NR	8.9±1.9, range NR	74±NR, range NR	NR	NR
Leonard et al. (1997) [73]	USA	Cross-sectional	29 families: 29 children, 29 mothers, 29 fathers	55% (16)	11.1±NR, range 8-14	0% (0) (28/29)	100% (29)	NR	NR	NR	3.8±NR, range 0.7-9.3
Leonard et al. (1998) [74]	USA	Cross-sectional	104 families: 104 parents (100% mothers)	NR	NR±NR, range NR, target group 8-17	NR	NR	NR	NR	NR	NR
Mansfield et al. (2004) [75]	USA	Cross-sectional	104 families: 104 children, 104 parents (sex parent NR)	53% (55)	11.7±NR, range NR, target group 8-17	7% (7)	NR	NR	NR	NR	2.7±1.5, range NR
Marker et al. (2018) [76]	USA	Cross-sectional	135 families: 133 children, 135 parents (84% mothers)	55% (73) (133/135)	13.5±1.8, range 10-16, target group 10-16 (133/135)	13% (18)	73% (99)	9.2±2.2, range NR	77±NR, range NR	87.4 (118)	5.6±3.4, range NR (130/135)
Marrero et al. (1995) [77]	USA	Intervention	106 families: 106 children	59% (63)	13.3±4.7, range NR, target group ≥5	4% (4)	NR	9.7±1.8, range NR	83±NR, range NR	0% (0)	6.2±4.5, range NR
Moreland et al. (2004) [79]	USA	Cross-sectional	153 families: 153 children, 153 parents (100% mothers)	44% (67)	12.9±2.3, range 8-16	NR	NR	8.4±1.4, range 5.9-14.3	68±NR, range 41-133	23% (35)	6.3±3.5, range 0.8-14.3
Mulvaney et al. (2013) [80]	USA	Cross-sectional	61 families: 61 adolescents	53% (32)	13.9±1.3, range NR, target group 12-17	15% (9)	NR	8.8±1.5, range NR	73±NR, range NR	49% (30)	5.8±3.3, range NR

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
			72 families: 72 parents (sex parent NR)	56% (40)	15.1±1.4, range NR, target group 13-17	10% (7)	NR	8.8±1.7, range NR	73±NR, range NR	58% (42)	6.3±3.4, range NR
Murphy et al. (2007) [82]	UK	Intervention	46 families (out of total of 78): 46 children, unclear if one parent or both parents completed responsibility measure (sex parent NR)	53% (41)	12.5±2.4, range NR, target group 6-16 (78/46)	NR	NR	9.1±1.2, range NR (78/46)	76±NR, range NR (78/46)	Unclear	4.7±3.2, range NR (78/46)
Murphy et al. (2012) [81]	UK	Intervention	306 families: Unclear who filled in the responsibility measure	48% (146)	13.1±1.9, range NR, target group 11-16	8% (23)	65% (197)	9.3±1.9, range NR	78±NR, range NR	7% (20)	5.6±3.3, range NR
Naar-King et al. (2007) [83]	USA	Intervention	127 families: 127 children, 127 parents ^k (88% mothers)	49% (62)	13.2±2.0, range NR, target group 10-17	73% (88) (121/127)	54% (69)	11.3±2.3, range NR	100±NR, range NR	8% (10)	4.9±3.1, range 1-13
Nansel et al. (2009) [85] ^A	USA	Cross-sectional	122 families: 122 children, unclear if one parent or both parents completed responsibility measure (sex parent NR)	50% (61)	12±NR, range NR, target group 9.0-14.5	28% (34)	NR	NR	NR	NR	NR
Neumark-Sztainer et al. (2002) [86]	USA	Cross-sectional	126 families (out of total of 143): 126 adolescents	51% (73) (143/126)	15.3±2.3, range NR, target group 12-21 (143/126)	10% (14) (143/126)	84% (120) (143/126)	8.8±1.6, range NR (143/126)	73±NR, range NR (143/126)	NR	NR
Ott et al. (2000) [87]	USA	Cross-sectional	143 families: 143 children, unclear if one parent or both parents completed responsibility measure (sex parent NR)	44% (63)	14.0±1.8, range 11-18	19% (27)	NR	NR	NR	0% (0)	5.6±3.7, range 1.1-15 years

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
Siminerio et al. (1999) [93]	USA	Intervention	32 families: 32 parents (if both parents were present, parents were asked to work as a team) (sex parent NR)	59% (19)	10.2±NR, range NR, target group 6-18	16% (5)	97% (31)	NR	NR	NR	NA
Streisand et al. (2005) [95]	USA	Cross-sectional	134 families: 134 parents ^k (86% mothers)	52% (70)	13.0±2.0, range NR, target group 9-17	NR	84% (113)	8.5±1.6, range 5.8-14.0	69±NR, range 40-130	20% (27)	4.9±3.20, range 0.5-14
Telo et al. (2015) [96]	USA	Cross-sectional	358 families: 358 children, 358 parents (sex parent NR)	50% (178)	13.0±2.8, range NR, target group 8-17.9	7% (25)	87% (310)	8.2±1.0, range NR	66±11, range NR	75% (268)	6.3±3.4, range NR
Vesco et al. (2010) [98] ^B	USA	Cross-sectional	261 families: 261 children, 261 parents ^k (82% mothers)	54% (141)	15.7±1.4, range NR, target group 13-18	13% (34)	78% (204)	9.0±1.8, range NR	75±NR, range NR	56% (146)	7.0±3.9, range 1-16.8
Weissberg-Benchell et al. (1997) [99]	USA	Cross-sectional	117 families: unclear if one parent or both parents completed responsibility measure (gender NR)	54% (63)	11.5±3.4, range 5-18	16% (19)	NR	7.6±1.2, range 5.0-11.9	60±NR, range 31-107	NR	5.6±3.1, range 1.1-16.3
Weissberg-Benchell et al. (2009) [101] ^A	USA	Cross-sectional	121 families: 121 children, 121 parents ^k (92% mothers)	50% (61)	12.1±1.6, range NR, target group 9-14.5	28% (34)	91% (110)	8.1±1.1, range NR	65±NR, range NR	32% (39)	5.4±3.2, range NR
Wilksch et al. (2013) [105]	AU	Intervention	20 families: unclear if one parent or both parents completed responsibility measure (sex parent NR)	0% (0)	11.1±0.6, range NR, target group 10-12 ^m	NR	NR	8.8±0.9, range NR ^m	73±NR, range NR ^m	30% (6)	4.2±2.9, range 1-10.1
Wu et al. (2013) [106] ^B	USA	Cross-sectional	133 families: 133 children, 133 parents (sex parent NR)	49% (65)	15.3±1.4, range 13.1-18.5 ^m	12% (16)	77% (103)	8.7±1.8, range NR ^m	72±20, range NR ^m	64% (85) ^m	5.9±3.8, range NR ^m

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
Wysocki et al. (1996) [109]	USA	Cross-sectional	100 families: 100 parents (sex parent NR)	41% (41)	12.3±3.1, range NR, target group 5-18	14% (14)	89% (89)	11.1±2.1, range NR	98±NR, range NR		5.2±3.1, range NR
Wysocki et al. (2009) [108]	USA	Cross-sectional	309 families: 309 parents ^k (87% mothers or stepmothers)	50% (155)	12.5±3.4, range NR, target group 9-14.5	24% (234)	NR	8.5±1.5, range NR	69±NR, range NR	42% (130)	5.4±1.5, range NR
<i>Subscale Responsibility of the Diabetes Responsibility and Conflict Scale (DRCS)</i>											
Berg et al. (2013) [38] ^E	USA	Cross-sectional	180 families: 180 adolescents, 176 mothers, 139 fathers	46% (82)	12.9±1.5, range 10.5-15.6	5% (9)	84% (148) (176/180)	NR	NR	57% (103)	4.5±2.9, range NR
Beveridge et al. (2006) [39] ^F	USA	Cross-sectional	127 families: 127 children, 127 parents (100% mothers)	52% (66)	12.9±1.7, range NR, target group 10-15	NR	86% (109)	NR	NR	NR	4.5±2.9, range NR
Butner et al. (2009) [41] ^E	USA	Cross-sectional	185 families: 185 children 185 mothers, 145 fathers	NR	12.5±1.3, Range NR, target group 10-14	6% (11.1)	NR	8.3±1.5, range 4.9-13.9	67±NR, range 30-128	NR	4.8±3.0, range NR
Cameron et al. (2007) [43]	NZ	Cross-sectional	47 families: 47 children, 47 parents (100% mothers)	45% (21)	15.9±1.1, range 13.2-17.9	17% (39)	77% (36)	8.3±1.3, range 4.7-14.0	67±NR, range 28-130	NR	7.3±4.1, range 1.0-14.6
Fortenberry et al. (2014) [6] ^E	USA	Longitudinal	213 families: 213 children ⁿ	NR	13.0±1.5, range NR, target group 10-14 ⁿ	6% (13)	NR	8.4±1.6, range NR ⁿ	68±NR, range NR ⁿ	51% (128) ^o (252/213)	NR
Grey et al. (2011) [54]	USA	Intervention study	129 families: 129 parents (sex parent NR)	41% (53)	8.0±2.9, range NR, target group 1-12	NR	NR	NR	NR	71% (92)	NR
Hanna et al. (2005) [56] ^G	USA	Cross-sectional	100 families: 100 children, 100 parents (80% mothers)	60% (60)	14.8±1.6, range 12.1-18.5	8% (8)	81% (81)	9.0±1.6, range 5.6-14.1	75±NR, range 38-131	NR	5.6±3.4, range: 1-15

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
Hanna et al. (2007) [57] ^G	USA	Cross-sectional	100 families: 100 children, 100 parents (80% mothers)	60% (60)	14.8±1.6, range NR, target group 12-18	9% (9)	81 (81)	9.0±1.6, range NR	75±NR, range NR	25% (25)	5.6±3.4, range NR
Hanna et al. (2013) [55] ^G	USA	Cross-sectional	87 families: 87 children, 87 parents ^k (81% mothers)	60% (52)	14.5±1.4, range NR, target group 12-17	9.0% (8) (86/87)	67% (58)	9.1±1.6, range NR	76±NR, range NR	26% (22) (86/87)	5.2±3.3, range NR
King et al. (2012) [69] ^E	USA	Longitudinal	252 families: 252 adolescents ^m	46% (115)	12.5±1.5, range 10-14 ^m	7.2% (18)	80% (201)	NR	NR	51% (129) ^m	4.7±3.0, range 1-12 ^m
King et al. (2014) [68] ^E	USA	Longitudinal	252 families: 252 mothers, 188 fathers ^m	46% (117)	12.5±1.5, range NR, target group 10-14 ^m	8.3% (21)	NR	NR	NR	51% (128) ^m	4.1±3.0, range NR ^m
Lee et al. (2015) [71]	Taiwan	Cross-sectional	210 families: 210 parents (sex parent NR)	47% (99)	14.0±2.2, range NR, target group 10-18	NR	92% (193)	8.5±1.9, range NR	69±NR, range NR	NR	6.6±3.9, range NR
Palmer et al. (2004) [89] ^F	USA	Cross-sectional	125 families: 125 children, 125 parents (100% mothers)	52% (66)	12.9±1.7, range NR, target group 10-15	NR	86% (108)	9.0±1.30, range 5.8-11.9	75±NR, range 40-107	NR	4.6±2.9, range NR
Palmer et al. (2009) [88] ^E	USA	Cross-sectional	185 families: 185 children, 185 mothers and 145 fathers	47% (87)	12.5±1.5, range 10.0-14.9 (145/185)	6% (11)	NR	8.1±1.3, range 4.9-13.9 (145/185)	65±NR, range 30-128 (145/185)	50% (92)	4.6±2.9, range 1-12 (n=145/185)
Srinivasan et al. (2004) [94]	AU	Intervention	110 families: unclear if one parent or both parents completed responsibility measure (sex parent NR)	NR	NR±NR, range 1.1-16.2	NR	NR	NR	NR	NR	NR

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
Turner et al. (2018) [97] ^E	USA	Longitudinal	199 families: 199 adolescents	46% (91)	12.4±1.5, range 10-14 ^m	6% (12)	NR	8.2±1.5, range NR ^m	66±NR, range NR ^m	56.2 (112) ^m	4.5±3.0, range NR ^m
Wiebe et al. (2005) [102] ^F	USA	Cross-sectional	127 families: 127 adolescents, 127 parents (100% mother)	52% (66)	12.8±1.7, range NR, target group 10-15	NR	86% (109)	8.9±1.5, range NR (95/127)	74±NR, range NR (95/127)	5% (6)	4.5±2.9, range NR
Wiebe et al. (2010) [104] ^E	USA	Cross-sectional	128 families: 128 children, 128 mothers, number of fathers unclear	57.0%	12.4±1.6, range 10.0-14.9	4% (5)	NR	7.9±1.2, range 4.9-13.6	63±NR, range 30-125	100% (128)	4.6±3.0, range 1.0-12.0
Wiebe et al. (2011) [4] ^F	USA	Longitudinal	82 families: 82 children, 82 parents (100% mothers)	52% (43)	12.8±1.7, range 10.0-16.0 ^m	1% (1)	90% (74)	8.9±1.3, range NR ^m (65/82)	74±NR, range NR ^m (65/82)	NR	NR
Wiebe et al. (2014) [103] ^E	USA	Longitudinal	252 families: 252 children, 252 mothers and 188 fathers	46% (117)	12.5±1.5, range NR, target group 10-14 ^m	8% (21)	NR	NR	NR	51% (128) ^m	4.1±3.0, range NR ^m
<i>CSII-Use survey</i>											
Rohan et al. (2015) [92] ^{H, q}	USA	Longitudinal	130 families: 130 parents (100% mothers) ^m	46% (109) (239/130)	10.5±0.9, range NR, target group 9-11 ^m (239/130)	25% (60) (239/130)	79% (188) ^m (239/130)	8.2±1.4, range 6-17 ^m (239/130)	66±NR, range 42-162 ^m (239/130)	100% (130)	4.4±2.5, range NR (239/130)
Weissberg-Benchell et al. (2007) [100]	USA	Cross-sectional	Unclear: 163 parents, unclear if one parent or both parents completed responsibility measure	47% (72)	10.6±3.8, range 3-18	4% (6)	92% (141)	NR	NR	100% (153)	NR
Wu et al. (2014) [107] ^{H, q}	USA	Longitudinal	130 families: 130 parents (100% mothers) ^m	46% (109) (239/130)	10.5±NR, range NR, target group 9-11 ^m (239/130)	25% (60) (239/130)	79% (188) ^m (239/130)	8.2±1.4, range NR ^m (239/130)	66±NR, range NR ^m (239/130)	100% (130)	4.4±2.5, range NR ^m (239/130)

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
<i>Deciding About Diabetes Treatment Scale (DADT)</i>											
Lewandowski et al. (2007) [13]	USA	Cross-sectional	51 families: 51 children, 51 parents (100% mothers)	47% (24)	14.67±1.2, range 13-18	12% (6)	71% (36)	8.9±1.95, range NR	74±NR, range NR	NR	6.3±4.2, range NR
Miller et al. (2003) [14]	USA	Cross-sectional	82 families: 82 children, 82 parents (100% mothers)	55% (45)	13.9±1.9, range 11-17	7% (6) (81/82)	NR	10.5±2.3, range 6.6-19.1	91±NR, range 49-185	NR	5.6±3.2, range 1-12.5
<i>Diabetes Regimen Responsibility Scale (DRRS)</i>											
Carey et al. (1997) [44]	USA	Cross-sectional	49 families: 49 children, 49 parents (94% mothers)	47% (23)	13.0±2.9, range 8-18	6% (3)	NR	8.6±2.0, range 4.8-13.0	70±NR, range 29-119	NR	5.7±2.7, range NR
Reid et al. (1994) [90]	USA	Cross-sectional	56 families: 56 parents (93% mothers)	50% (28)	13.03±3.0, range 8-18	5% (2)	80% (45)	8.3±2.1, range 4.8-13.0	67±NR, range 29-119	NR	5.4±2.7, range 1-11.3
<i>Adapted version of the Diabetes Independence Survey (DIS) – child version</i>											
Botello-Harbaum et al. (2008) [40]	USA	Cross-sectional	81 families: 81 children	44% (36)	13.3±1.7, range 11-16	15% (12)	NR	NR	NR	63% (51)	7.7±3.7, range 5.2-14.9
<i>Adapted version of the Diabetes Independence Survey (DIS) – parent version</i>											
Rohan et al. (2015) [92] ^{H, q}	USA	Longitudinal	109 families: 109 parents (100% mothers) ^m	46% (109)	10.5±0.9, range NR, target group 9-11 ^m (239/109)	25% (60) (239/109)	79% (188) ^m (239/109)	8.2±1.4, range 6-17 ^m (239/109)	66±NR, range 42-162 ^m (239/109)	0% (0)	4.4±2.5, range NR (239/109)
Wu et al. (2014) [107] ^{H, q}	USA	Longitudinal	109 families: 109 parents (100% mothers) ^m	46% (109)	10.5±NR, range NR, target group 9-11 ^m (239/109)	25% (60) (239/109)	79% (188) (239/109)	8.2±1.4, range NR ^m (239/109)	66±NR, range NR (239/109)	0% (0)	4.4±2.5, range NR ^m (239/109)
<i>Independent functioning in daily and nondaily diabetes management checklist</i>											
Hanna et al. (2003)a [58] ^l	USA	Cross-sectional	34 families: 34 adolescents	44% (15)	14.7±2.0, range NR, target group 11-18	12% (4)	NR	8.8±NR, range NR	73±NR, range NR	0% (0)	7.5±3.8, range NR

Study ^b	Country	Design ^c	Number of unique families: valid n completing ^d (mothers)	Sex child ^e (male)	Age child ^f (years)	Ethnicity ^e (Non-White ^g)	Household composition ^e (Two-parent household ^h)	HbA _{1c} ⁱ (%) ^f	HbA _{1c} ^j (mmol/mol) ^f	Treatment (CSII) ^e	Diabetes duration (years) ^f
Hanna et al. (2003) ^b [59] ^l	USA	Cross-sectional	27 families: 27 adolescents	48% (13)	15.0±1.9, range 11-18	11% (3)	63% (17)	8.5±1.3, range NR	69±NR, range NR	NR	7.0±3.3, range NR
<i>Children's Diabetes Inventory (CDI)</i>											
McNabb et al. (1994) [78]	USA	Intervention	24 families: unclear if one parent or both parents completed responsibility measure (sex parent NR)	54% (13)	9.9±NR, range NR, target group 8-12 ^m	NR	NR	11.8±3.6, range NR ^m	105±NR, range NR	NR	NR
<i>Nameless instrument described by Friedemann-Sanchez et al. 2018</i>											
Friedemann-Sanchez et al. (2018) [50]	India	Cross-sectional	178 families: unclear if one parent or both parents completed responsibility measure (50% mothers)	47% (83)	12.3±0.3, range NR, target group 0-18	NR	NR	10.6±NR, range NR	92±NR, range NR	0 (0)	4.7±NR, range NR
<i>Nameless instrument described by Nakamura et al. 2010</i>											
Nakamura et al. (2010) [84]	Japan	Cross-sectional	368 families: unclear whether child or parent completed responsibility measure (sex parent NR)	38% (138)	14.0±2.6, range NR, target group 9-18	NR	NR	8.0±1.5, range 4.9-14.9	64±NR, range 30-139	NR	6.5±3.9, range 0.3-18
<i>One question regarding insulin administration described by Gowers et al. 1995</i>											
Gowers et al. (1995) [52]	UK	Cross-sectional	60 families: unclear whether child or parent completed responsibility measure (sex parent NR)	60% (47)	13.9±1.2, range NR, target group 12-16	NR	NR	10.1±1.93, range NR	87±NR, range NR	0% (0)	NR
<i>Combination of two previously described questionnaires (DIS and CSII-Use survey) described by Robinson et al. 2011</i>											
Robinson et al. (2011) [91]	USA	Longitudinal	86 families: 86 children, 86 parents ^k (87% mothers, 11% fathers, 2% other)	45% (39)	12.9±2.0, range NR, target group 10-16 ^m	26% (22)	78% (67)	NR	NR	26% (22) ^m	6.3±3.8, range NR ^m

Abbreviations: AU= Australia; CSII= Continuous Subcutaneous Insulin Infusion; FU= Follow-up; HbA_{1c}= Glycated haemoglobin; NR= Not reported; NZ= New Zealand; UK= United Kingdom.

^a The notation (n/n) indicates that the number of respondents that provided data on the responsibility measure differs from the displayed n: (n of displayed data/n that filled in the questionnaire).

^b Data displayed as name first author (year of publication).

^c Design according to the assessment of the responsibility measure.

^d Number of unique families of whom information regarding the division of diabetes care responsibility was obtained: valid n of family members who completed responsibility measures.

^e Data displayed as %, n (percentages were round up to whole numbers and might therefore not add up to 100).

^f Data displayed as mean±SD (round up to one decimal), if available range and else target group (i.e., targeted group as described in the in- and exclusion criteria).

^g Non-white: non-white (and non-Hispanic), Non-Caucasian, Non-European American, minority (in European and American studies).

^h Two-parent household: Intact families, both biological parents living together, number of parents married (if both parents are included data of mothers is displayed), living with both biological parents or living with one biological parent and one stepparent.

ⁱ HbA_{1c} or GhB value as reported in Table 1, includes most recent HbA_{1c} measure, average HbA_{1c} of the past year, average rating across three measurements (testing, 3 months and 6 months before testing), or average HbA_{1c} in the year following the assessment.

^j If only HbA_{1c} values in % were reported, HbA_{1c} values in mmol/mol were computed (formula mmol/mol= (10.93 x %) -23).

^k Primary caregiver (parents who is most involved in the diabetes care of the child) was included.

^l Countries within Europe, Japan, Australia and North America.

^m Baseline characteristics in case of assessments on multiple time points.

ⁿ At time point 2 of a larger study.

^o At time point 1 of a larger study.

^p Standardized value.

^q Mothers of children using pen therapy completed the DIS; Mothers of children using pump therapy completed the CSII-Use survey.

A,B,C,D,E,F,G,H,I= indicates which studies are most likely to be based on the same sample.

Table 3. Designs of instruments measuring the division of diabetes care responsibilities between parents and children according to the original article¹ and the use of these instruments across the included studies (n= 84).

Instrument name – original study ^a	Mode of administration	Target population ^b	Description ^c	Response options ^c	Original language (country) ^c	Notes regarding the use of the scale across the included studies ^d
Diabetes Family Responsibility Questionnaire (DFRQ) – Anderson et al. 1990 [8]	Questionnaire	<ul style="list-style-type: none"> • Children aged 6 to 21 • Parents of children aged 6 to 21 	For 17 individual diabetes care tasks responders rate who takes responsibility or initiates responsibility for the particular task. For parental ratings three subscales were identified: general health maintenance (7 items), regimen tasks (6 items), social presentation (4 items)	<ul style="list-style-type: none"> • “1= Parent take or initiate responsibility for this almost all of the time” • “2= Parent and child share responsibility for this about equally” • “3= Child takes or initiates responsibility for this almost all of the time” 	English (USA)	<p>Some studies included a subset of items or revised items (6 items from the subscale regimen tasks, 10, 15 or 19 items) [10, 45, 47, 81, 83, 95]. With respect to instructions, in one study responders were asked to indicate which family member had the most responsibility for each task.[47] Additionally, response options were further specified (“father”, “mother”) or new options were added (“no one takes responsibility”, “not applicable”, “someone else”, “no one in the home”) [5, 47, 63, 83].</p> <p>In 11 studies, items were reversely scored (“1= Child responsibility”, “3= Parent responsibility”) [9, 12, 37, 46, 62, 67, 75, 79, 82, 96, 98]; for 27 studies value labels were not explicitly stated on an item level [5, 10, 36, 42, 47, 48, 51, 53, 60, 63-66, 70, 72, 74, 77, 80, 81, 83, 85, 86, 99, 101, 106, 108, 109].</p> <p>(n= 49/n= 42)</p>
Responsibility subscale of the Diabetes Responsibility and Conflict Scale	Questionnaire	<ul style="list-style-type: none"> • Children aged 10 to 15 • Parents of children aged 10 to 15 	Besides the degree of conflict between parents and children for 15 diabetes related activities, responders also indicate the division of tasks within the family per activity.	Unclear from original study. However, included studies consistently described a 5-point scale with the answer options [4, 6, 38, 39, 41, 43,	English (USA)	In some studies, items were revised, updated, edited and deleted, resulting in versions with 12, 14, 23, 24 and 25 items [4, 6, 38, 39,

Instrument name – original study ^a	Mode of administration	Target population ^b	Description ^c	Response options ^c	Original language (country) ^c	Notes regarding the use of the scale across the included studies ^d
(DRCS) –Rubin et al. 1989 [110] ^e			Based on two studies that named original items, tasks included performing (“who tests urine”, “who talks to adults about your diabetes”) and decision making (“who decides which people know that you are diabetic”, “who remembers to make appointments with dentists or other doctors”) [43, 71].	55-57, 71, 88, 89, 97, 102-104: <ul style="list-style-type: none"> • Child does the task alone • Task is shared with parents • Parent does the task alone 		41, 43, 68, 71, 88, 89, 94, 97, 102-104]. In one study the instructions of this scale were revised to incorporate specific questions per treatment modality (pen/pump) [55]. Value labels differed between studies; most studies labelled response options from 1 to 5 [4, 6, 38, 39, 41, 43, 71, 88, 89, 97, 102-104] or from 0-4 [56, 57], with higher values indicating more parent responsibility (“0/1= Child”; “2/3= Shared”; “4/5= Parent”). In one study, value labels were reversed for children (“0= Parent” to “4= Child”) [57]. Additionally, in one study higher items scores reflected more shared responsibility (“2= Shared”, “0= Parent”, “0= Child”, 1= responses in between) [55]. For 4 studies, response labels were not explicitly stated [54, 68, 69, 94]. In a few studies, the instrument was completed by families of children up to 16 [94] and 18 years old [43, 55-57, 71] or by families of children from 1 year old [54, 94]. (n= 20/n= 7)

Instrument name – original study ^a	Mode of administration	Target population ^b	Description ^c	Response options ^c	Original language (country) ^c	Notes regarding the use of the scale across the included studies ^d
Responsibility subscales of the CSII-Use survey – Weissberg-Benchell et al. 2007 [100]	Questionnaire	Parents of children aged 3 to 18	<p>Next to assessing the child’s mastery level for 23 relevant skills for pump use, parents are also asked to indicate per task who is responsible for ensuring that the task is completed.</p> <p>This scale has four subscales including CSII operations (e.g., prepares infusion site properly) (12 items), knowledge of diabetes management (e.g. knows causes of low/high blood sugars) (12 items), comfort in social situations (e.g., talks to peers about diabetes) (3 items), monitoring of blood sugar (e.g., checks blood sugars 4-6 times per day).</p>	<ul style="list-style-type: none"> • Parent is responsible • Child is responsible • Responsibility is shared 	English (USA)	A modified version of 28-items was used for pump users in 2 articles which were based on the same sample [92, 107]. Additionally, in both studies, the following value labels were assigned: “1= Caregiver”, “2= Shared”, “3= Child” [92, 107]. (n= 3/n= 2)
Deciding About Diabetes Treatment Scale (DADT) – Saletsky, 1991 [112] ^f	Questionnaire	<ul style="list-style-type: none"> • Children aged 11 to 18 • Parents of children aged 11 to 18 	Unclear from the original study. The two studies that included this instrument both described that for every diabetes self-care task, responders indicated the perceived level of autonomy the child has in making decisions about diabetes care tasks [13, 14].	<p>Unclear from original study. However, the two included studies both described a 4-point scale with the answer options [13, 14]:</p> <ul style="list-style-type: none"> • Parent tells the child to complete the tasks/the mother takes exclusive responsibility for decisions about the task • Parent and child discuss the task, but the parent has the final say about it/makes the final decision • The parent and child discuss the task, but the child has the final say about it/the child makes the final decision 	English (USA)	As the reference study provided limited information, it is difficult to say whether scales were used in the same way as in the reference study. Between the two studies [13, 14], small differences were observed regarding the wording of the response options. In one study, the number of items and exact response values were described; the scale consisted of 20 items (e.g., how much to eat at meals, when to test blood, how much insulin to take) and response values ranged from 1 to 4, with higher scores indicating more child responsibility for decision making [13]. In the other study, the number of items and exact values of response options were not explicitly stated [14]. No other differences regarding the

Instrument name – original study ^a	Mode of administration	Target population ^b	Description ^c	Response options ^c	Original language (country) ^c	Notes regarding the use of the scale across the included studies ^d
				<ul style="list-style-type: none"> The child decides on his or her own to complete the task/the child assumes complete responsibility for the task 		<p>administration of the scale were observed [13, 14].</p> <p>(n= 2/n= 2)</p>
Diabetes Regimen Responsibility Scale (DRRS) - Ruggiero et al. 1991 [111] ^g	Unclear from the original study. The two included studies that used this instrument both describe a questionnaire [44, 90]	Unclear from the original study. The two included studies that used this instrument both children aged 8 to 18 [44] and parents of children aged 8 to 18 completed the questionnaire [44, 90]	Unclear from the original study. The two included studies that used this instrument both described that responders report the division of diabetes care separately for 42 individual diabetes care tasks covering six self-care areas (i.e., insulin administration, dose adjustment, glucose testing, ketone testing, dietary management, regimen modifications) [44, 90].	Unclear from original study	Unclear from the original study. The two included studies that used this instrument were both conducted in English (USA) and did not describe that the scales had to be translated [44, 90]	<p>As the original article was not accessible, it is unclear to what extent the scales have been used as described in the original article. In one study only the subscales fingerpicks (5 tasks) and diet (6 items) were included [90].</p> <p>Information regarding instructions and response options differed between the two included studies. One study described that responders were instructed to indicate who assumes responsibility (exact number and content of response options and recall period were not described) [44] whereas in the other study responders were requested to specify the responsibility children had for treatment tasks during the previous month on a 4 point scale (range “1= 0-25%” to “4= 76-100%”) [90].</p> <p>(n=2/n=2)</p>
Adapted version of the Diabetes Independence Survey (DIS) –	Questionnaire	<ul style="list-style-type: none"> Children aged 11-16 	The DIS was originally developed to examine the extent of the child’s mastery (e.g., independent performing without requiring physical or verbal assistance to	<p>In both adapted versions three answer options were reported:</p> <ul style="list-style-type: none"> Parent 	English (USA)	In a study that was based on the same sample as the original study of the adapted DIS for parents, the

Instrument name – original study ^a	Mode of administration	Target population ^b	Description ^c	Response options ^c	Original language (country) ^c	Notes regarding the use of the scale across the included studies ^d
child version – Botello-Harbaum et al. 2008 [40]; parent version – Wu et al. 2014 [107]		<ul style="list-style-type: none"> Parents of children aged 9-11 	<p>complete it correctly) as reported by parents for 38 diabetes care skills, regarding management of hypo- and hyperglycemia, blood glucose testing, insulin injections, diet and exercise. Listed skills were related to performing tasks (“draws”, “recognizes”, “rotates”, “performs”), knowledge about diabetes and treatment (“states”, describes”) and cognitive skills like planning and anticipation within diabetes care. In two studies, authors adapted instructions and response options to examine the division of diabetes care responsibilities within families.</p> <p>In the first adapted version that was developed for children, a modified subset of the original set of tasks was used (29 instead of 38 items) and responders were instructed to indicate who took responsibility for diabetes care tasks [40].</p> <p>In the second adapted version that was developed for parents of children using a pen to administer insulin, original items were updated in order to reflect modern diabetes management and parents were asked to specify who in the family was primarily responsible for monitoring and completing the tasks [107].</p>	<ul style="list-style-type: none"> Shared Child/Youth <p>In the version for children, values from 1 to 3 were assigned with higher scores reflecting more parental responsibility.</p> <p>In the version for parents, values from 1 to 3 were assigned with higher scores reflecting more child responsibility.</p>		<p>use of the scale did not differ from the earlier provided description.</p> <p>(child: n= 1/n= 1; parent: n= 2/n= 1)</p>
Independent functioning in daily and nondaily diabetes management checklist – Hanna et al. 2003 [58]	Questionnaire	Children aged 11 to 18	This questionnaire consists of two sets of tasks (adapted from the DFRQ): diabetes care tasks which are performed daily (e.g., glucose monitoring, insulin administration, food choices, exercise and dental hygiene) (14 items) and non-daily (i.e., keeping track of supplies, making and keeping	<ul style="list-style-type: none"> “2= Adolescent” “1= Parent” “1= Both adolescent and parent” “1= No one” 	English (USA)	<p>In a later study, answers value labels ranged from 0 (“Parent”, “Both adolescent and parent”, “No one”) to 1 (“Child”).</p> <p>(n= 2/n= 1)</p>

Instrument name – original study ^a	Mode of administration	Target population ^b	Description ^c	Response options ^c	Original language (country) ^c	Notes regarding the use of the scale across the included studies ^d
			appointments, discussing health concerns with diabetes care team, paying for health expenses, watching weight, foot care and discussing diabetes with adults and friends) (20 items). For all tasks responders indicate who performed the task. In order to facilitate recall, for daily tasks adolescents are asked to report who performed the task the previous day whereas for non-daily tasks adolescents indicated who performed the task the previous three months.			
Children’s Diabetes Inventory (CDI) – McNabb et al. 1994 [78]	Questionnaire	Parents of children aged 8 to 12	Next to the rating of behavioral frequency for each of the 35 diabetes self-care behaviors (e.g., insulin administration, meal planning, making adjustments in the diabetes regimen communicating with others about diabetes (care), caring for diabetes equipment and supplies), the degree to which the child assumed responsibility for performing these behaviors was reported.	5-point scale ranging from answer categories: “1= Parent totally responsible” to “5= Child totally responsible”	English (USA)	Not applicable as this measure is only included in the instrument’s reference. (n= 1)
Friedemann-Sanchez et al. 2018 [50]	Questionnaire	Parents of children under 18	Responders identify for each direct and indirect care task the primary person who carried out or oversaw this task. Direct tasks included determining insulin dose and injection site, drawing-up and injecting insulin, checking for daytime/nocturnal hypoglycemia, measuring and recording blood glucose levels, determining the timing and content of meal. Indirect tasks included calling the clinic for insulin dose adjustments or emergencies, taking the child to clinic/hospital, monitoring diabetes supplies, keeping insulin cold, monitoring glucose trends, and communicating with teachers about diabetes.	<ul style="list-style-type: none"> • Child • Mother • Father • Other 	India (Hindi or Marathi)	Not applicable as this measure is only included in the instrument’s reference. (n= 1)

Instrument name – original study^a	Mode of administration	Target population^b	Description^c	Response options^c	Original language (country)^c	Notes regarding the use of the scale across the included studies^d
Nakamura et al. 2010 [84]	Questionnaire	Children with type 1 and type 2 diabetes aged 9 to 22	This scale assesses family involvement and contains 37 items regarding insulin injection (5 items), self-monitoring blood glucose (4 items), management of hypoglycemia (5 items), diet (8 items), exercise (5 items), contact with school (4 items) and outpatient visits (6 items).	“1= Yes”	Japanese (Japan)	Not applicable as this measure is only included in the instrument’s reference. (n= 1)
Gowers et al. 1995 [52]	Question(naire)	Parents of children aged 12 to 16	Parents were asked about their involvement in the diabetes regimen – the only item that is described is the administration of insulin injections. Unclear if the scale consists of more items.	<ul style="list-style-type: none"> Child gives injections alone Parent and child give injections together Unclear if more answer options were included	English (UK)	Not applicable as this measure is only included in the instrument’s reference. (n= 1)
Robinson et al. 2003 [91]	Questionnaire	<ul style="list-style-type: none"> Children aged 10 to 16 Parents of children aged 10 to 16 	This scale (40 items) is a combination of two included questionnaires that assess the division of diabetes care responsibilities (Diabetes Independence Survey and the Diabetes Behavior Rating Scale). Responders indicate per task whose job it is in the family to see that the task is done.	5-point scale ranging from answer categories: <ul style="list-style-type: none"> It’s all my job It’s all my (parent’s /child’s) job 	English (USA)	Not applicable as this measure is only included in the instrument’s reference. (n= 1)

Abbreviations: CSII= Continuous Subcutaneous Insulin Infusion; UK= United Kingdom.

^a Article that describes the development of the scale or the first article that introduces the instrument (i.e., no information is available regarding the development of the instrument).

^b If information regarding the development of the scale is available, target population is described. If information regarding the development of the scale is not available, the population in the first article that introduces the instrument is described. If the original study could not be accessed, information was supplemented with consistently reported information in the included studies.

^c Extracted from the original study. If the original study could not be accessed or provided very limited information, information was supplemented with consistently reported information in the included studies.

^d n/n= total number of included studies using the instrument/number of studies based on unique samples using the instrument.

^e Conference abstract – in 2007, an updated version focusing solely on conflict was published in 2007 [117].

^f Thesis – only abstract available.

^g Article not available.

Table 4. Psychometric properties of the included instruments (n= 13).

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
Diabetes Family Responsibility Questionnaire (DFRQ) - Anderson et al. 1990 [8]	<p>1a. Aims to measure perceptions of who takes responsibility for a broad range of diabetes tasks that contribute to the child's diabetes management and, consequently, to examine disagreement and agreement regarding perceptions of responsibility between parents and children</p> <p>1c. Children (6-21 years), mothers and health care providers were represented in the elicitation of items</p> <p>2a. Targeted to children (6-21 years) and their parents</p> <p>2b. Children (6-21 years) and mothers were represented in the elicitation of items</p> <p>2c. Concept elicitation was based on interviews, resulting in 22 items</p> <p>3. Scale was developed to use in research and clinical practice in order to work with family systems around adherence and support issues</p> <p><i>Reviewer comments</i></p> <p>1a. Unclear how authors define responsibility</p> <p>1b. Origin of construct unclear</p> <p>1c. Unclear to what extent items refer to the construct of interest according to concept elicitation, literature or professionals</p> <p>2b. Fathers were not involved, although the scale is targeted to parents</p> <p>2c. Details of concept elicitation are not</p>	<p><i>Anderson et al., (1990)[8]: n mothers and children (6-20 years) from the same family= 121 – factor analysis performed as part of the development of the questionnaire</i></p> <p>1a. <i>Methodological information</i></p> <ul style="list-style-type: none"> • Not explicitly stated if factor analysis is based on a reflective model • Unclear if unidimensionality or structural validity is of concern • Explanatory factor analysis was conducted • n= 121 for original 22 items, n per item= 5.5; not meeting criteria of n>200 • Data inspection techniques not reported • Factor analytic method was principal component analysis for parent and child data separately; PCA is not a true form of factor analysis • Factor retention method not mentioned • Criteria for item factor selection included high loadings on more than one factor or failing to load on a meaningful factor; definitions of high loadings and loadings on more than factor are unclear • Oblique rotation method; not stated if it was expected that factors were correlated, turned out that factors were correlated <p>1b. <i>Results</i> <i>Parent data (mothers n= 121)</i></p>	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> • For one study it was unclear whether the child or parent completed the responsibility measure (α= NR) [81], for one study it was unclear if children, next to parents, also completed the responsibility measure [45], for three studies it was unclear for which responder internal consistency was reported (α range= 0.74-0.88 [51, 53, 72]) • For parent reports, 22 studies based on unique samples examined internal consistency of the total scale [8, 10-12, 36, 42, 47-49, 61, 64, 66, 67, 70, 75, 83, 95, 98, 99, 101, 105, 108], of which three studies also reported internal consistencies of subscales [8, 75, 98]; 14 studies did not include internal consistency analysis [5, 37, 45, 60, 63, 73, 74, 76, 79, 82, 87, 93, 96, 109] • For child reports, 18 studies based on unique samples examined internal consistency of the total scale [8, 10-12, 36, 42, 47, 48, 61, 64, 66, 67, 70, 75, 83, 86, 98, 101], of which two 	<p>3a. Concurrent validity</p> <p>i. <i>Methodological information</i></p> <ul style="list-style-type: none"> • Maternal (sub)scale scores were correlated with maternal ratings of the independence subscale of the Family Environment Scale (FES, Moos, 1986), measuring independence as a priority for individual family members; measurement properties of the comparator instrument are not reported; only focussed on maternal reports <p>ii. <i>Results</i></p> <ul style="list-style-type: none"> • Maternal total scores and scores on the subscale Regimen Task correlated with the Family Environment Scale subscale "independence" ($r= 0.27$, $p<0.01$; $r= 0.21$, $p<0.05$), mothers with higher reports of independence as a priority for individual family members reported greater child responsibility for regimen tasks and total tasks of the DFRQ; Unclear to which extent hypotheses

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
	<p>provided – unclear how and to what extent families were involved</p> <p>5. Not applicable, as no recall period is defined</p> <p>6. No comprehensiveness study conducted; does not include tasks related to new technologies</p> <p>9. Responsibility might be differently interpreted by responders (behavioral performance vs. “feeling” responsibility); shared responsibility can refer to frequency (i.e. parent half of the time, child half of the time) or to shared performance/decisions</p> <p>10. Assumes that all tasks are relevant within families; no responsibility is not necessarily a result of discrepant views about who is responsible for tasks</p>	<ul style="list-style-type: none"> For parents, 17 of 22 items loaded on three factors (five items were dropped): General health maintenance (eigenvalue= 5.10, 39% of variances explained), Regimen tasks (eigenvalue= 1.85, 11% of variances explained), Social presentation of diabetes (eigenvalue= 1.54, 9% of variances explained); no performance of a visual scree plot analysis or parallel analysis, Celicer’s MAP Factor loadings: <ul style="list-style-type: none"> General health maintenance: 0.43-0.77; one item >0.3 on other factor; factor load difference for all items is >0.20 on all alternative factors; one item loaded on more than one factor (>0.30), however, difference in item load was more than 0.20 Regimen tasks: 0.41-0.73; 1 item >0.3 on other factor; factor load difference is <0.20 for 1 item on one alternative factor; 1 item loaded on more than 1 factor (>0.30 other factor) and the item load difference was <0.20 Social presentation of diabetes: 0.48-0.76; no other items loading more than 0.3 on other factors; factor load difference for all items is >0.20 on all alternative factors Oblique rotation: r general health maintenance-regimen tasks= 0.41, $p<0.01$; r general health maintenance – social presentation= 0.26, $p<0.01$; r regimen tasks-social presentation= 0.21, $p>0.01$; r general health maintenance – total scale= 0.85, $p<0.01$; r regimen tasks – 	<p>studies also reported internal consistencies of subscales [75, 98]; 12 studies did not include internal consistency analysis [5, 37, 60, 63, 73, 76, 77, 79, 80, 82, 87, 96]</p> <p>2b. <i>Results</i> <i>Parents</i></p> <ul style="list-style-type: none"> Total scale: α ranged from 0.64 (fathers)-0.92 (22/36) [8, 10-12, 36, 42, 47-49, 61, 64, 66, 67, 70, 75, 83, 95, 98, 99, 101, 105, 108] Number of studies $\alpha<0.70$ total scale: two; [64, 66] both studies had a small sample size ($n= 30$ [66] and $n= 45$ [64]); in one study only paternal scores were examined [64]– was not evaluated in other studies) Subscales Anderson et al. 1990: α general health maintenance= 0.73-0.78 [8, 75]; α Regimen tasks= 0.63-0.79 [8, 75]; α Social presentation of diabetes= 0.59-0.79 [8, 75] Number of studies $\alpha<0.70$ subscales Anderson et al. 1990: 1 – only for Regimen Tasks and Social presentation of diabetes [75]; could be a time effect (2004 vs. 1990) or this could be 	<p>were in line with the results as hypotheses regarding associations with (sub)scales and the direction and magnitude of the effect were not predefined; comparator instrument measures a different construct.</p>

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
		<p>total scale= 0.83, $p < 0.01$; r social presentation – total scale= 0.62, $p < 0.01$</p> <p><i>Child data</i> (n= 121)</p> <ul style="list-style-type: none"> The structure of parent data did not emerge in child reported data nor from adolescents specifically <p><i>Vesco et al., (2010) [98] – 261 families of children 13-18 years old – both parent and child report – factor analysis performed on the final instrument</i></p> <p>1a. <i>Methodological information</i></p> <ul style="list-style-type: none"> Not explicitly stated if factor analysis is based on a reflective model Analyses were executed in order to evaluate structural validity – hypotheses are based on findings regarding diabetes-specific conflicts and state that both child and parent reports load on 2 factors (indirect and direct diabetes care tasks); states that previous factor structure as reported by Anderson et al. 1990 was only valid for parental reports of children >12 years, not in line with results as reported by Anderson et al. 1990 n= 261 for 17 items, n per item= 15.4 Data inspection techniques not reported Reports CFA for both child and parent reports; used EFA techniques Interpretation of the factor structure was based on the scree plot, eigenvalues, and the rotated factor pattern which produced standardized 	<p>a result of poor structural validity or a small number of items</p> <ul style="list-style-type: none"> Subscales Vesco et al. 2010: α direct= 0.80 [98]; α indirect= 0.49 [98] Number of studies $\alpha < 0.70$ subscales Vesco et al. 2010: 1 – only for indirect tasks [98]; this could be a result of poor structural validity or a small number of items <p><i>Children</i></p> <ul style="list-style-type: none"> Total scale: α ranged from 0.51-0.85 (18/30) [8, 10-12, 36, 42, 47, 48, 61, 64, 66, 67, 70, 75, 83, 86, 98, 101] Number of studies $\alpha < 0.70$ total scale: 3 [47, 66, 101]; 2 studies had a small sample size (n= 31 [47], n= 30 [66]), in one study instructions and response options were modified [47] Subscales Anderson et al. 1990: α general health maintenance= 0.58 [75]; α Regimen tasks= 0.60 [75]; α Social presentation of diabetes= 0.43 [75] Number of studies $\alpha < 0.70$ subscales Anderson et al. 1990: 1 (all subscales) [75]; structural validity was not established for 	

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
		<p>regression coefficients; unclear how methods were used to determine the number of underlying factors</p> <ul style="list-style-type: none"> • Varimax rotation method; does not allow factors to be correlated, not stated by authors if this was also their expectation <p><i>1b. Results</i> <i>Parent data (n= 261)</i></p> <ul style="list-style-type: none"> • For parents, 11 items loaded on the factor direct tasks (eigenvalue= 3.42) whereas 6 items loaded on the indirect tasks (eigenvalue= 0.82); no display of a visual scree plot or performance of parallel analysis, Celicer's MAP • Factor loadings: <ul style="list-style-type: none"> ○ Direct tasks: 0.34-0.64; 4 items had an item load of <0.4 on the primary factor; no other items loading more than 0.3 on other the factor; for one item the factor load difference was >0.20 on the other factor; four items had low factor loadings (<0.4), for one item the load difference was >0.20, however, the factor loading on the other factor was less than 0.3 ○ Indirect tasks: 0.28-0.50; four items had an item load of <0.4 on the primary factor; no other items loading more than 0.3 on other the factor; factor load difference >0.20 on other factor for two items; four items had low factor loadings (<0.4), for two items the load difference was >0.20, however, the 	<p>child report, could be a time effect (2004 vs. 1990) or could be caused by a small number of items</p> <ul style="list-style-type: none"> • Subscales Vesco et al. 2010: α direct= 0.73 [98]; α indirect= 0.57 [98] • Number of studies α<0.70 subscales Vesco et al. 2010: 1 (only indirect tasks); this could be a result of poor structural validity or a small number of items 	

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
		<p>factor loadings on the other factor were less than 0.3</p> <ul style="list-style-type: none"> • Varimax rotation method: <i>r</i> indirect and direct tasks= 0.35 <p><i>Child data (n= 261)</i></p> <ul style="list-style-type: none"> • As for the parent data, for the child data the same 11 items loaded on the factor direct tasks (eigenvalue= 2.74) whereas six items loaded on the indirect tasks (eigenvalue= 0.94); no display of a visual scree plot or performance of parallel analysis, Celicer's MAP • Factor loadings: <ul style="list-style-type: none"> ○ Direct tasks: 0.30-0.54; four items had an item load of <0.4 on the primary factor; two items had factor loadings >0.3 on other the factor; factor load difference >0.20 on other factor for four items; four items had low factor loadings (<0.4), two of these items had a load difference of <0.2, however, the factor loading on the other factor was less than 0.3; two additional items had a factor loading of more than 0.3 on the other factor, in both cases the factor load difference was less than 0.2 between the two factors ○ Indirect tasks: 0.31-0.52; two items had an item load of <0.4 on the primary factor; none of the items had a factor loading of >0.3 on other the factor; factor load difference >0.20 on other factor for one item; two items had low factor loadings (<0.4), one of these items had a load 		

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
		<p>difference of <0.2 with the other factor however, the factor on the other factor was less than 0.3</p> <ul style="list-style-type: none"> • Varimax rotation method: <i>r</i> indirect and direct tasks= 0.33 <p>Other reported subscales– no factor analysis reported [60]</p>		
<p>Part of the Responsibility of the Diabetes Responsibility and Conflict Scale (DRCS) - Rubin et al. 1989 [110]^b</p>	<p>1a. Aims to measure the division of responsibility and degree of conflict between parents and children for diabetes related activities 2a. Targeted to children (10-15 years) and their parents</p> <p><i>Reviewer comments</i> 1a. Unclear how authors define responsibility 1b. Origin of construct unclear 1c. Unclear to what extent items refer to the construct of interest according to concept elicitation, literature or professionals 2. Unclear to what extent items are relevant for the target population of interest 3. Context of use unclear 5. Not applicable, as no recall period is defined 6. Comprehensiveness study not reported; does not include tasks related to new technologies 9. Responsibility might be differently interpreted by responders (behavioral performance vs. “feeling” responsibility);</p>	Not reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> • For parent reports, five studies based on unique samples examined internal consistency of the total scale [4, 43, 54, 56, 71] and one study only reported internal consistencies for general items and pump specific items separately [104]; one study did not include internal consistency analysis [94] • For child reports, four studies based on unique samples examined internal consistency of the total scale [4, 56, 97, 104] and one additional study based on the same sample as one of the studies examining the total scale reported internal consistencies for general items and pump specific items [104] <p>2b. <i>Results</i> <i>Parents</i></p>	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
	shared responsibility can refer to frequency (i.e. parent half of the time, child half of the time) or to shared performance/decisions 10. Assumes that all tasks are relevant within families; no response option to indicate that tasks are not relevant within the family		<ul style="list-style-type: none"> Total scale: α ranged from 0.81-0.91 (5/7) [4, 43, 54, 56, 71] Subscales: α general items= 0.89-0.91; α pump specific= 0.88-0.91 [104] <p><i>Children</i></p> <ul style="list-style-type: none"> Total scale: α ranged from 0.82-0.92 (4/4) [4, 43, 56, 97] Subscales: α general items= 0.90; α pump specific= 0.87 [104] 	
Responsibility subscales of the CSII-Use survey –Weissberg-Benchell et al. 2007 [100]	<p>1a. Aims to assess CSII-specific skill mastery and autonomy (responsibility that for ensuring that each particular tasks was completed)</p> <p>1c. The questionnaire was developed by interviewing nurse educators who were certified CSII trainers and reviewing CSII care manuals to capture CSII-related tasks</p> <p>2a. Targeted to parents of children aged 3 to 18 years</p> <p>3. Clinical purposes</p> <p>8a. Items were then reviewed by a small group of parents and diabetes clinicians for clarity and accuracy. Revisions were made based on feedback to create the final version.</p> <p><i>Reviewer comments</i></p> <p>1b. Origin of the construct unclear</p>	Explanatory factor analysis performed as part of the development of the questionnaire, but it is unclear whether the analysis was based on mastery- or responsibility scores [100]	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> Of the two studies based on unique samples [92, 100], one study reported internal consistency [92], the other study did not include an internal consistency analysis [100] <p>2b. <i>Results – parents</i></p> <ul style="list-style-type: none"> Total scale: range α across three years= 0.88-0.91 (1/2) [92] 	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
	<p>2b. Not reported that the target population was represented in elicitation of relevant items</p> <p>4. No justification of response options</p> <p>5. Not applicable as no recall period was defined</p> <p>6. Not reported that a comprehensiveness study was conducted</p> <p>7. Not reported that a comprehensibility study regarding instructions was conducted</p> <p>8b. Results and methodological information of the comprehensibility study unclear, final items were not tested</p> <p>9. Difficult to judge whether questions are appropriately worded as exact items are not displayed for responsibility</p> <p>10. Assumes that all tasks are relevant within families; no response option to indicate that tasks are not relevant within the family</p>			
Deciding About Diabetes Treatment Scale (DADT) – Saletsky, 1991 [112] ^c	Unclear as the original study could not be assessed	Not reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> For parent reports, two studies based on unique samples examined internal consistency of the total scale [13, 14] For child reports, two studies based on unique samples examined internal consistency of the total scale [13, 14] <p>2b. <i>Results Parents</i></p>	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
			<ul style="list-style-type: none"> Total scale: $\alpha = 0.93-0.95$ (2/2) [13, 14] <p><i>Children</i></p> <ul style="list-style-type: none"> Total scale: $\alpha = 0.87-0.93$ (2/2) [13, 14] 	
Diabetes Regimen Responsibility Scale (DRRS) - Ruggiero et al. 1991 [111] ^d	Unclear as the original study could not be assessed	Subscales defined in subsequent studies [44, 90] – no factor analysis reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> For parents, two studies based on unique samples included an internal consistency analysis [44, 90]; one reported internal consistencies for the two included subscales [90] whereas the other study reported the internal consistency of the total scale and the mean and range of subscales [44] Only one study included child reports; for child reports internal consistency of the total scale and the mean and rang of subscales was reported [44] <p>2b. <i>Results</i></p> <p><i>Parents</i></p> <ul style="list-style-type: none"> Total scale: $\alpha = 0.89$ (1/1) [44] Subscales: Fingerpricks $\alpha = 0.87$ [90]; Diet $\alpha = 0.73$ [90]; mean and range of 6 subscales $\alpha = 0.79$, range 0.67-0.89 [44] (2/2); for at least one subscale an insufficient internal consistency was 	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
			<p>reported, this could be a result of poor structural validity or a small number of items</p> <p><i>Children</i></p> <ul style="list-style-type: none"> • Total scale: $\alpha = 0.88$ (1/1) [44] • Mean and range of 6 subscales: $\alpha = 0.66$; range 0.56-0.83 (1/1) [44]; for at least one subscale an insufficient internal consistency was reported, this could be a result of poor structural validity or a small number of items 	
Adapted version of the Diabetes Independence Survey (DIS) – child version– Botello-Harbaum et al. 2008 [40]	<p>1a. Aims to assess parent and child responsibility for diabetes management activities</p> <p>1c. Selection of key diabetes care tasks was adapted from the DIS [113, 114]</p> <p>2a. Targeted to parents of children from 9 to 11 years old</p> <p>4. Response options were adapted from the DFRQ [8]</p> <p><i>Reviewer comments</i></p> <p>1a. Unclear how authors define responsibility</p> <p>1b. Origin of construct unclear</p> <p>1c. DIS was originally developed to measure mastery of self-care skills, this construct is conceptually different from the division of diabetes care responsibilities. Unclear to what extent items refer to the construct of</p>	Not reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> • Only the original study used this questionnaire and examined internal consistency of the total scale [40] <p>2b. <i>Results – children</i></p> <ul style="list-style-type: none"> • Total scale: $\alpha = 0.83$ (1/1) [40] 	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
	<p>interest according to concept elicitation, literature or professionals</p> <p>2b. Target group was not reported to be involved in the election of relevant items</p> <p>3. Context of use unclear</p> <p>4. No justification of response options in study that is referred to</p> <p>5. Not applicable, as no recall period is defined</p> <p>6. No comprehensiveness study conducted; not reported that scale does include tasks related to new technologies</p> <p>9. Responsibility might be differently interpreted by responders (behavioral performance vs. “feeling” responsibility); shared responsibility can refer to frequency (i.e., parent half of the time, child half of the time) or to shared performance/decisions</p> <p>10. Assumes that all tasks are relevant within families; no responsibility is not necessarily a result of discrepant views about who is responsible for tasks</p>			
Adapted version of the Diabetes Independence Survey (DIS) – parent version – Wu et al. 2014 [107]	<p>1a. Aims to measure who resided responsibility for the diabetes regimen</p> <p>1c. Items of the DIS [113, 114] were adapted (e.g., items to reflect modern diabetes management recommendations such as rechecking blood glucose sugars when readings are high)</p> <p>2a. Targeted to parents of 9 to 16 year old children</p> <p>4. Response options were adapted from the CSII-use survey [100]</p>	Not reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> Only the original study and a study based on the same sample used this questionnaire and examined the internal consistency of the total scale [92, 107] <p>2b. <i>Results – parents</i></p>	

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
	<p><i>Reviewer comments</i></p> <p>1a. Unclear how authors define responsibility</p> <p>1b. Origin of construct unclear</p> <p>1c. DIS [113, 114] was originally developed to measure mastery of self-care skills, this construct is conceptually different from the division of diabetes care responsibilities. Unclear to what extent items refer to the construct of interest according to concept elicitation, literature or professionals</p> <p>2a. Only targeted to children using an insulin pen</p> <p>2b. Target group was not reported to be involved in the election of relevant items</p> <p>3. Context of use unclear</p> <p>4. No justification of response options in study that is referred to</p> <p>5. Not applicable, as no recall period is defined</p> <p>6. No comprehensiveness study conducted</p> <p>9. Responsibility might be differently interpreted by responders (behavioral performance vs. “feeling” responsibility); shared responsibility can refer to frequency (i.e., parent half of the time, child half of the time) or to shared performance/decisions</p> <p>10. Assumes that all tasks are relevant within families; no responsibility is not necessarily a result of discrepant views about who is responsible for tasks</p>		<ul style="list-style-type: none"> Total scale: range α across three years $\alpha = 0.94-0.95$ (1/1) [92] 	

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
Independent functioning in daily and nondaily diabetes management checklist - Hanna et al. 2003 [58]	<p>1a. Aims to measure independent functioning in daily and nondaily diabetes management</p> <p>1c. Items were adapted from the DFRQ [8], items were added in an effort to capture the universe of diabetes management activities related to glucose monitoring, insulin administration, diet, exercise, and general health behaviors. A physician researcher who was an expert in adolescence and diabetes reviewed the diabetes management tasks items for thoroughness and appropriateness</p> <p>2a. Targeted to children (early, middle and late adolescence – 11 to 18 years)</p> <p>5. A recall period of the previous day was included to facilitate recall of daily diabetes care tasks</p> <p><i>Reviewer comments</i></p> <p>1b. Origin of construct unclear</p> <p>1c. No information about to what extent items refer to the construct according to concept elicitation, literature or professionals</p> <p>2b. Not reported that target population was involved in the development</p> <p>3. Context of interest unclear</p> <p>4. No justification for the response options</p> <p>6. No comprehensiveness study conducted; unclear whether tasks are also related to new technologies</p>	Not reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> The two studies that were based on the same sample reported internal consistencies, however, one study [58] reported separate internal consistencies for daily and non-daily functioning whereas the other study [59] only reported internal consistency for the total scale Kuder-Richardson alpha reliability was calculated <p>2b. <i>Results - children</i></p> <ul style="list-style-type: none"> Daily functioning: Kuder-Richardson= 0.80 (1/1) [58] Non-daily functioning: Kuder-Richardson= 0.80 (1/1) [58] Total scale: Kuder-Richardson= 0.87 (1/1) [59] 	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
	9. Difficult to judge whether questions are appropriately worded as exact items are not displayed			
Children's Diabetes Inventory (CDI) - McNabb et al. 1994 [78]	<p>1a. Aims to measure the degree to which the child assumed responsibility for performing diabetes self-care behaviors</p> <p>1c. Prior to administration, the questionnaire was reviewed by a panel of two pediatric diabetologists, two diabetes nurse specialists, one diabetes nutrition specialist and two health educators</p> <p>2a. Parents of children aged 8 to 12</p> <p>3. Was developed for research purposes</p> <p><i>Reviewer comments</i></p> <p>1a. Unclear how authors define responsibility</p> <p>1b. Origin of construct unclear</p> <p>1c. No information about to what extent items refer to the construct according to concept elicitation, literature or professionals</p> <p>2b. not reported that target population was involved in the development</p> <p>4. No justification for the response options</p> <p>5. Not applicable as no recall period was defined</p> <p>6. No comprehensiveness study conducted; unclear whether tasks are also related to new technologies</p> <p>9. Difficult to judge whether questions are appropriately worded as exact items are not displayed</p>	Subscales defined by the original study [78] – no factor analysis reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> Only the original study used this questionnaire and examined internal consistency on a total scale [78] <p>2b. <i>Results - parents</i></p> <ul style="list-style-type: none"> Total scale: $\alpha = 0.94$ (1/1) [78] 	<p><i>Comparison between subgroups (known-groups validity)</i></p> <p>3bi. <i>Methodological information</i></p> <ul style="list-style-type: none"> Age groups were contrasted (8-10 years vs. 11 to 12 years) – statistical methods unclear; characteristics of age groups are unclear, unclear if t-test was used Correlation analysis was performed between age and the responsibility measure. <p>3bii. <i>Results</i></p> <ul style="list-style-type: none"> Hypotheses regarding the comparison between subgroups and the relation between age and responsibility were not predefined Younger children (8-10 years) were reported by parents to be given significantly less overall responsibility for their diabetes self-care than were older children (11-12 years) ($p < 0.01$); as hypotheses

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
				<p>were not predefined, it is unclear to what extent results are in line with hypotheses</p> <ul style="list-style-type: none"> The reported level of children's overall responsibility for diabetes self-care was significantly correlated with age ($r = .64$; $p < 0.01$), with younger children reported as taking less responsibility for their diabetes self-care than older children; as hypotheses were not predefined, it is unclear to what extent results are in line with hypotheses
Friedemann-Sanchez et al. 2018 [50]	<p>1a. Aims to identify the primary person performing each diabetes care tasks 2a. Parents of children under 18 years</p> <p><i>Reviewer comments</i> 1b. Origin of construct unclear 1c. No information about to what extent items refer to the construct according to concept elicitation, literature or professionals 2b. Not reported that target population was involved in the development 3. Context of interest unclear 4. No justification for the response options 5. Not applicable as no recall period was defined</p>	Subscales defined by the original study [50] – no factor analysis reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> Only the original study used this questionnaire but did not report information regarding internal consistency (0/1) [50] 	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
	6. No comprehensiveness study conducted; unclear whether tasks are also related to new technologies 9. Difficult to judge whether questions are appropriately worded as exact items are not displayed			
Nakamura et al. 2010 [84]	2a. Targeted to children with type 1 and type 2 diabetes (aged 9 to 22) <i>Reviewer comments</i> 1a. Description of the construct of interest unclear 1b. Origin of construct unclear 1c. No information about to what extent items refer to the construct according to concept elicitation, literature or professionals 2b. Not reported that target population was involved in the development 3. Context of interest unclear 4. No justification for the response options 5. Not applicable as no recall period was defined 6. No comprehensiveness study conducted; unclear whether tasks are also related to new technologies 9. Difficult to judge whether questions are appropriately worded as exact items are not displayed	Subscales described by the original study [84] – no factor analysis reported	2a. <i>Methodological information</i> • Only the original study used this questionnaire and examined internal consistency on the total scale [84] 2b. <i>Results - children</i> • Total scale: $\alpha = 0.82$ (1/1) [84]	Not reported
Gowers et al. 1995 [52]	Unclear as the development study provides very limited information	Not reported	2a. <i>Methodological information</i> • Only the original study used this question(naire) but did not report	Not reported

Instrument name – original study ^a	Content validity (provided information by the original study and reviewers' comments)	Internal structure		Other measurement properties (i.e., inter-rater reliability, hypothesis test for construct validity) (available information from the original study; comments)
		Structural validity (available information from the included studies; comments)	Internal consistency (available information from the included studies; comments)	
			information regarding internal consistency (0/1) [52]	
Robinson et al. 2011 [91]	<p>1a. Aims to assess the extent of parent responsibility</p> <p>1c. Items were based on items of the DIS [113, 114] and the CDI [78]</p> <p>2a. Targeted to children with type 1 or type 2 diabetes (aged 9 to 22)</p> <p><i>Reviewer comments</i></p> <p>1a. Unclear how authors define responsibility</p> <p>1b. Origin of construct unclear</p> <p>1c. No information about to what extent items refer to the construct according to concept elicitation, literature or professionals</p> <p>2b. Not reported that target population was involved in the development</p> <p>3. Context of interest unclear</p> <p>4. No justification for the response options</p> <p>5. Not applicable as no recall period was defined</p> <p>6. No comprehensiveness study conducted; unclear whether tasks are also related to new technologies</p> <p>9. Difficult to judge whether questions are appropriately worded as exact items are not displayed</p>	Not reported	<p>2a. <i>Methodological information</i></p> <ul style="list-style-type: none"> Only the original study used this questionnaire and examined internal consistency on a total scale [91] <p>2b. <i>Results</i></p> <p><i>Parents</i></p> <ul style="list-style-type: none"> Total scale: $\alpha = 0.93$ (1/1) [91] <p><i>Children</i></p> <ul style="list-style-type: none"> Total scale: $\alpha = 0.93$ (1/1) [91] 	Not reported

Abbreviations: CSII= Continuous Subcutaneous Insulin Infusion.

^a Article that describes the development of the scale or the first article that introduces the instrument (i.e., no information is available regarding the development of the instrument).

^b Conference abstract - In 2007, an updated version focusing solely on conflict was published in 2007 [117].

^c Thesis –only abstract available.

^d Article not available.

Table 5. Overall rating of psychometric properties (R)^a and the quality of evidence (Q)^b per instrument based on the COSMIN guideline (n= 13).

	DFRQ [8]		DRCS [110]		CSII-Use survey [100]		DADT [112]		DRRS [111]		DIS – child [40]		DIS – parent [107]		Checklist Hanna et al. [58]		CDI [78]		Friedmann -Sanchez et al. [50]		Nakamura et al. [84]		Gowers et al. [52]		Robinson et al. 2011 [91]		
	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	R	Q	
Content validity ^c		L		?		VL		?		?		VL		VL		VL		VL		VL		VL		?		?	VL
<i>Relevance</i>	+		?		-		?		?		-		-		-		-		-		-		?		-		
<i>Comprehensiveness</i>	-		?		?		?		?		?		?		?		?		?		?		?		?		?
<i>Comprehensibility</i>	±		?		?		?		?		?		?		?		?		?		?		?		?		?
Structural validity ^d	?	M	*		*		*		*		*		*		*		*		*		*		*		*		*
Internal consistency ^e	?	M	?	M	?	L	?	M	?	VL	?	VL	?	L	?	VL	?	VL	?	L	?	L	*		?	VL	
Construct validity	?	L	*		*		*		*		*		*		*		?	VL	*		*		*		*		*

Abbreviations: CDI= Children’s Diabetes Inventory; CSII= Continuous Subcutaneous Insulin Infusion; DADT= Deciding About Diabetes Treatment scale; DFRQ= Diabetes Family Responsibility Questionnaire; DIS= Diabetes Independence Survey; DRCS= Diabetes Responsibility and Conflict Scale; DRRS= Diabetes Regimen Responsibility Scale.

^a Rating of psychometric properties (R): + = sufficient; - = insufficient; ? = indeterminate; ± = inconsistent; * = structural validity, internal consistency and/or construct validity were not examined.

^b Quality of evidence (Q): High (H), moderate (M), low (L), very low (VL).

^c For some instruments aspects of content validity were scored as “?” as the development study did not provide sufficient information and the instrument was not available.

^d In line with the COSMIN guideline, only studies testing existing instruments are included, not studies refining or developing instruments.

^e Within the COMIN guideline, at least low evidence for sufficient structural validity is required to score internal consistency as sufficient.

DISCUSSION

This systematic review identified 13 instruments (all questionnaires) measuring the division of diabetes care responsibilities, which have been used in 84 studies representing 62 unique samples. Across studies, the DFRQ [8] was by far most frequently used; most other instruments were only used in two studies or less [40, 50, 52, 58, 78, 84, 91, 100, 107, 111, 112]. For most instruments, key information regarding psychometric properties was missing or showed suboptimal quality and there was a wide variety in scoring methods. Furthermore, clear construct definitions and links with theory were lacking.

Based on the findings from this review, there are several aspects that warrant specific attention when selecting an instrument to assess the division of diabetes care responsibilities. Firstly, across questionnaires, variations were observed with respect to the target group, recall period, instruction, the number and content of tasks and the included response options. In order to select the scale that best fits the objectives of users, these variations in general features should be taken into account. Secondly, the psychometric properties of the instrument should be considered. Most instruments received comparable poor ratings on most criteria for content validity. Yet, the DFRQ [8] received a higher rating for content validity in the current analysis based on the involvement of the target population in the development of the scale. In contrast to the other identified scales, for the DFRQ also structural validity was examined. Nevertheless, the structural validity was evaluated to be indeterminate due to insufficient reporting and methodological flaws. While taking into account that sufficient structural validity is lacking for all questionnaires and it was unclear whether reflective or formative models were of concern, internal consistencies were overall sufficient. Construct validity was only examined for the DFRQ [8] and CDI [78], however, both comparisons had methodological flaws. Thirdly, feasibility could affect the choice for a certain scale to assess the division of diabetes care responsibilities. For five scales, including the DFRQ, both parent and child versions are available [8, 91, 110-112] and only the DFRQ and DRCS have been translated into other languages [8, 110]. None of the included studies reported the instrument's completion time or possible costs. Finally, it should be considered whether addition of measurement of constructs related to the division of diabetes care responsibilities is appropriate for the objectives of users. Currently scores are evaluated by sample distributions as no standards are available indicating the "right" division of diabetes care responsibilities. Since the use of sample distributions limits comparability among studies, several instruments could be added to examine the experiences with the division of tasks, such as the evaluation of the division [44, 118], appraisal of parental involvement [10, 119, 120], the level of experienced conflict around the division of diabetes care tasks [74], the extent of received advice from others regarding the division [74], decision making about the division [58] or parental expectations of age appropriate levels of responsibility [44, 73]. Furthermore, all instruments were orientated to some extent towards behavioral aspects of responsibility. Although scales that use "responsibility" in their instructions do not rule out cognitive/emotional aspects of responsibility (e.g., feeling responsible), cognitive aspects of the transfer of diabetes care responsibilities could be measured by questionnaires targeting the child's and parent's readiness to shift the balance of responsibility [64], parental reasons to transfer responsibilities [89] and perceived advantages and disadvantages of the transfer of diabetes care responsibilities of children and parents [57]. Additionally, during this systematic search two instruments were excluded from the analysis as a broader spectrum of constructs was assessed [121, 122]. One of these instruments is the interview developed by Anderson et al., examining the level of parental involvement (i.e., both responsibility and

monitoring) in blood glucose monitoring and insulin injections [121]. Compared to questionnaires, the use of an interview could provide more detailed information.

There are also several aspects to consider while scoring and interpreting scales. First of all, some studies generated a family score by averaging parent- and child-reports [60, 61, 63, 91]. Although similar results have been obtained when analyses were conducted with separate parent and child responsibility scores compared to aggregated parent and child scores [61], this method is not preferable as a broad correlation range has been reported between parent- and child-reports (range $r = 0.36-0.75$) [4, 8, 10, 14, 42, 55, 61, 63, 70, 79, 88, 89, 91, 96, 101, 102, 106]. Second, given the results and methodological shortcomings regarding the evaluation of the internal structure of the included questionnaires, at this time, total scores are preferred and subscale scores should be interpreted with caution. Third, to obtain final (sub)scale scores, items can either be added or averaged across items, or the number/percentage of tasks per answer category can be calculated. As some authors suggest that shared responsibility is preferable throughout adolescence [5], shared responsibility is often targeted in studies. However, to examine this type of responsibility specifically, shared responsibility should be contrasted with parent and child responsibility at either an item level (i.e., defining value labels) or during scoring (i.e., calculating the number/percentage of tasks per answer category) in order to interpret scores correctly; if one end of the scale presents child responsibility and the other end parent responsibility, middle scores across items do not necessarily indicate shared responsibility as scores can be reported solely as the parent's or child's responsibility. Accordingly, a study that used both mean scores and the percentage of tasks per answer category (i.e., solely child responsibility, shared responsibility, solely parent responsibility) in the analyses showed that the two scoring techniques can give different results [5]. When mean scores were used, beneficial effects on self-efficacy were associated with more parental responsibility whereas in the analysis with the percentages of tasks per answer category beneficial results were associated with more shared responsibility and less parent responsibility [5]. When shared responsibility is not of special concern or scales do not include shared responsibility as a separate response option, mean scores are preferred over total scores in order to facilitate comparisons between studies; across studies items have been deleted because of project objectives or applicability.

Strengths of this study are the use of a systematic search and that studies which used the instruments were included alongside with the original study, which provided information regarding the actual use of instruments and allowed for a more thorough evaluation of the internal consistency. Moreover, the assessment of the quality of the psychometric properties was thoroughly evaluated by using standard checklists. However, several limitations should be taken into account. First of all, for most scales the evaluation of content validity by reviewers was restricted to example items and general descriptions of scales as full questionnaires were not available [40, 50, 52, 58, 78, 84, 91, 100, 107, 110-112]. Secondly, in the current review methodological quality was considered as low if original studies provided limited information regarding psychometric properties and/or did not report the development of the scale in detail. As six of the 13 identified instruments were developed before 2000 [8, 52, 78, 110-112] and at the time guidelines regarding questionnaire development were not introduced yet, it is unclear whether poor ratings for methodological quality are rather a result of reporting than compromised methodological quality. Thirdly, relevant articles could have been missed as the search was performed in two databases and studies published after 1st June 2018 were not evaluated for eligibility.

As this evaluation did not identify any instruments with high quality support for content validity, future studies are warranted. These studies could more thoroughly evaluate the psychometric properties of recent instruments, update instruments that do not contain tasks around new technologies in diabetes care (e.g., most recent versions of insulin pump, glucose monitoring devices) and evaluate the validity of this scale with parents and children, or develop a new questionnaire. During the development of a new scale or the update of an existing scale, guidelines such as the Food and Drug Administration guideline and COSMIN Stud Design Checklist should be closely followed in order to facilitate adequate psychometric properties [123, 124]. Finally, “responsibility” appears to be a concept that is difficult to define; instructions and answer options can be differently interpreted across respondents and might be cultural dependent. Therefore, comprehensibility and cross-cultural validity are especially important aspects to monitor to reduce measurement error.

CONCLUSION

Several instruments have been used to assess the division of diabetes care responsibilities. As there were large variations in general characteristics of instruments, researchers and clinicians should closely compare their objectives to the strengths and limitations of these instruments. At this point, the DFRQ might be most suitable to assess the division as this instrument facilitates comparability because of its frequent use and it underwent more thorough psychometric evaluation in comparison to the other instruments, although, with suboptimal results. In order to get a better understanding of the division and transfer of diabetes care responsibilities, the DFRQ needs to be updated to include contemporary diabetes tasks or a new scale should be developed that addresses the reported limitations of the DFRQ. Ideally, future scales should be linked to theory and revised or developed in line with methodological guidelines.

LIST OF ABBREVIATIONS

- CDI = Children’s Diabetes Inventory
- CSII = Continuous Subcutaneous Insulin Infusion
- DADT = Deciding About Diabetes Treatment Scale
- DFRQ = Diabetes Family Responsibility Questionnaire
- DIS = Diabetes Independence Survey
- DRCS = Diabetes Responsibility and Conflict Scale
- DRRS = Diabetes Regimen Responsibility Scale

AUTHORS’ CONTRIBUTION

JA, GN, EH and FP obtained funding and designed the study. LN, PW, EvM and HJA reviewed the search string. JA performed the systematic search and subsequently, JA and GN screened the results. JA extracted relevant information from the included studies, GN, EH, LN and FP checked these extractions. JA, GN, EH and FP drafted the manuscript LN, PW, EvM, and HJA reviewed the manuscript.

CONSENT FOR PUBLICATION

All authors approved the final version of this review for publication.

AVAILABILITY OF DATA AND MATERIALS

Not applicable

FUNDING

This study was supported by the Dutch Research Council (NWO) (Research Talent, 406-14-103). The funding source had no role in the design of the study, the analysis of data, the preparation of the manuscript and the decision to submit the manuscript.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare that are relevant to the content of this article.

REFERENCES

- [1] Scott LK. Developmental Mastery of Diabetes-Related Tasks in Children. *Nurs Clin North Am*. 2013;48(2):329-42.
- [2] Kyngäs H. Compliance of adolescents with chronic disease. *J Clin Nurs*. 2000;9(4):549-56.
- [3] Young MT, Lord JH, Patel NJ, Gruhn MA, Jaser SS. Good cop, bad cop: quality of parental involvement in type 1 diabetes management in youth. *Curr Diab Rep*. 2014;14(11):546.
- [4] Wiebe DJ, Gelfand D, Butler JM, Korbel C, Fortenberry KT, McCabe JE, et al. Longitudinal associations of maternal depressive symptoms, maternal involvement, and diabetes management across adolescence. *J Pediatr Psychol*. 2011;36(7):837-46.
- [5] Helgeson VS, Reynolds KA, Siminerio L, Escobar O, Becker D. Parent and adolescent distribution of responsibility for diabetes self-care: links to health outcomes. *J Pediatr Psychol*. 2008;33(5):497-508.
- [6] Fortenberry KT, Berg CA, King PS, Stump T, Butler JM, Pham PK, et al. Longitudinal trajectories of illness perceptions among adolescents with type 1 diabetes. *J Pediatr Psychol*. 2014;39(7):687-96.
- [7] Pierce JS, Aroian K, Schifano E, Milkes A, Schwindt T, Gannon A, et al. Health Care Transition for Young Adults With Type 1 Diabetes: Stakeholder Engagement for Defining Optimal Outcomes. *J Pediatr Psychol*. 2017;42(9):970-82.
- [8] Anderson BJ, Auslander WF, Jung KC, Miller JP, Santiago JV. Assessing family sharing of diabetes responsibilities. *J Pediatr Psychol*. 1990;15(4):477-92.
- [9] Anderson BJ, Holmbeck G, Iannotti RJ, McKay SV, Lochrie A, Volkening LK, et al. Dyadic measures of the parent-child relationship during the transition to adolescence and glycemic control in children with type 1 diabetes. *Fam Syst Health*. 2009;27(2):141-52.
- [10] Cameron FJ, Skinner TC, de Beaufort CE, Hoey H, Swift PG, Aanstoot H, et al. Are family factors universally related to metabolic outcomes in adolescents with Type 1 diabetes? *Diabet Med*. 2008;25(4):463-8.
- [11] Lancaster BM, Gadaire DM, Holman K, LeBlanc LA. Association Between Diabetes Treatment Adherence and Parent-Child Agreement Regarding Treatment Responsibilities. *Fam Syst Health*. 2015.
- [12] Law GU, Walsh J, Queralt V, Nouwen A. Adolescent and parent diabetes distress in type 1 diabetes: the role of self-efficacy, perceived consequences, family responsibility and adolescent-parent discrepancies. *J Psychosom Res*. 2013;74(4):334-9.
- [13] Lewandowski A, Drotar D. The relationship between parent-reported social support and adherence to medical treatment in families of adolescents with type 1 diabetes. *J Pediatr Psychol*. 2007;32(4):427-36.
- [14] Miller VA, Drotar D. Discrepancies between mother and adolescent perceptions of diabetes-related decision-making autonomy and their relationship to diabetes-related conflict and adherence to treatment. *J Pediatr Psychol*. 2003;28(4):265-74.
- [15] Austin S, Guay F, Senecal C, Fernet C, Nouwen A. Longitudinal testing of a dietary self-care motivational model in adolescents with diabetes. *J Psychosom Res*. 2013;75(2):153-9.
- [16] Beal SJ, Riddle IK, Kichler JC, Duncan A, Houchen A, Casnellie L, et al. The Associations of Chronic Condition Type and Individual Characteristics With Transition Readiness. *Acad Pediatr*. 2016;16(7):660-7.
- [17] Butner JE, Berg CA, Munion AK, Turner SL, Hughes-Lansing A, Winnick JB, et al. Coordination of Self- and Parental-Regulation Surrounding Type I Diabetes Management in Late Adolescence. *Ann Behav Med*. 2018;52(1):29-41.

- [18] Cemeroglu AP, Can A, Davis AT, Kleis L, Daniel MS, Rapp SM, et al. Comparison of the expectations of caregivers and children with type 1 diabetes mellitus for independence in diabetes care-related tasks. *Endocr Pract.* 2014;20(7):629-37.
- [19] Cheraghi F, Shamsaei F, Mortazavi SZ, Moghimbeigi A. The Effect of Family-centered Care on Management of Blood Glucose Levels in Adolescents with Diabetes. *Int J Community Based Nurs Midwifery.* 2015;3(3):177-86.
- [20] Dashiff C, Vance D, Abdullatif H, Wallander J. Parenting, autonomy and self-care of adolescents with Type 1 diabetes. *Child Care Health Dev.* 2009;35(1):79-88.
- [21] Doherty FM, Calam R, Sanders MR. Positive parenting program (Triple P) for families of adolescents with type 1 diabetes: A randomized controlled trial of self-directed teen Triple P. *J Pediatr Psychol.* 2013;38(8):846-58.
- [22] Ellis DA, Podolski C-L, Frey M, Naar-King S, Wang B, Moltz K. The role of parental monitoring in adolescent health outcomes: Impact on regimen adherence in youth with type 1 diabetes. *J Pediatr Psychol.* 2007;32(8):907-17.
- [23] Grover S, Bhadada S, Kate N, Sarkar S, Bhansali A, Avasthi A, et al. Coping and caregiving experience of parents of children and adolescents with type-1 diabetes: An exploratory study. *Perspect Clin Res.* 2016;7(1):32-9.
- [24] Markowitz JT, Volkening LK, Laffel LMB. Care utilization in a pediatric diabetes clinic: Cancellations, parental attendance, and mental health appointments. *J Pediatr.* 2014;164(6):1384-9.
- [25] Morrison S, Dashiff C, Abdullatif H, Moreland E. Parental separation anxiety and diabetes self-management of older adolescents: a pilot study. *Pediatr Nurs.* 2012;38(2):88-95.
- [26] Niba LL, Aulinger B, Mbacham WF, Parhofer KG. Predictors of glucose control in children and adolescents with type 1 diabetes: results of a cross-sectional study in Cameroon. *BMC Res Notes.* 2017;10(1):207.
- [27] Pate T, Klemencic S, Battelino T, Bratina N. Fear of hypoglycemia, anxiety, and subjective well-being in parents of children and adolescents with type 1 diabetes. *J Health Psychol.* 2016;24(2):209-18.
- [28] Pendley JS, Kasmien LJ, Miller DL, Donze J, Swenson C, Reeves G. Peer and family support in children and adolescents with type 1 diabetes. *J Pediatr Psychol.* 2002;27(5):429-38.
- [29] Rodriguez LM, Hassan KV, Rhodes J, McKay SV, Heptulla RA. The Relationship between Self-Management and Glycemic Control in a Cohort of Children with Type 1 and Type 2 Diabetes. *J Diabetes Metab.* 2013;4.
- [30] Shahpurwala Z. Youth with type 1 diabetes mellitus: An examination of adolescents' illness perceptions and associated impact on health-related quality of life. US: ProQuest Information & Learning; 2016.
- [31] Higgins JP, Green S. *Cochrane handbook for systematic reviews of interventions* Version 5.1.0 [updated March 2011]. 2011. The Cochrane Collaboration. Available from: www.handbook.cochrane.org
- [32] Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res.* 2018;27(5):1171-9.
- [33] Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147-57.
- [34] Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res.* 2018;27(5):1159-70.
- [35] Howard MC. A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? *International Journal of Human-Computer Interaction.* 2016;32(1):51-62.
- [36] Agarwal S, Jawad AF, Miller VA. A multivariate model exploring the predictive value of demographic, adolescent, and family factors on glycemic control in adolescents with type 1 diabetes. *Pediatr Diabetes.* 2016;17:500-8.
- [37] Anderson BJ, Vangsness L, Connell A, Butler D, Goebel-Fabbri A, Laffel LM. Family conflict, adherence, and glycaemic control in youth with short duration Type 1 diabetes. *Diabet Med.* 2002;19(8):635-42.
- [38] Berg CA, Butner JE, Butler JM, King PS, Hughes AE, Wiebe DJ. Parental persuasive strategies in the face of daily problems in adolescent type 1 diabetes management. *Health Psychol.* 2013;32(7):719-28.
- [39] Beveridge RM, C AB, D JW, D LP. Mother and adolescent representations of illness ownership and stressful events surrounding diabetes. *J Pediatr Psychol.* 2006;31(8):818-27.
- [40] Botello-Harbaum M, Nansel T, Haynie DL, Iannotti RJ, Simons-Morton B. Responsive parenting is associated with improved type 1 diabetes-related quality of life. *Child Care Health Dev.* 2008;34(5):675-81.
- [41] Butner J, Berg CA, Osborn P, Butler JM, Godri C, Fortenberry KT, et al. Parent-adolescent discrepancies in adolescents' competence and the balance of adolescent autonomy and adolescent and parent well-being in the context of Type 1 diabetes. *Dev Psychol.* 2009;45(3):835-49.

- [42] Caccavale LJ, Nansel TR, Quick V, Lipsky LM, Laffel LM, Mehta SN. Associations of disordered eating behavior with the family diabetes environment in adolescents with Type 1 diabetes. *J Dev Behav Pediatr*. 2015;36(1):8-13.
- [43] Cameron LD, Young MJ, Wiebe DJ. Maternal trait anxiety and diabetes control in adolescents with type 1 diabetes. *J Pediatr Psychol*. 2007;32(7):733-44.
- [44] Carey TC, Reid G, Ruggiero L, Horner J. The Diabetes Regimen Responsibility Scale: Information on internal consistency and validity in a pediatric sample. *Assessment*. 1997;4(2):207-9.
- [45] Cemeroglu AP, Timmer S, Turfe Z, Davis AT, Koehler TJ, Can A, et al. Differences in parental involvement in the care of children and adolescents with type 1 diabetes mellitus on multiple daily insulin injections versus continuous subcutaneous insulin infusion. *J Pediatr Endocrinol Metab*. 2016;29(3):265-72.
- [46] Cortina S, Repaske DR, Hood KK. Sociodemographic and psychosocial factors associated with continuous subcutaneous insulin infusion in adolescents with type 1 diabetes. *Pediatr Diabetes*. 2010;11(5):337-44.
- [47] Dashiff CJ. Self- and dependent-care responsibility of adolescents with IDDM and their parents. *J Fam Nurs*. 2003;9(2):166-83.
- [48] Delamater AM, Daigre AL, Marchante AN, Pulgarón ER, Patiño-Fernandez AM, Sanchez J. Intrinsic motivation in ethnic minority youth with type 1 diabetes. *Child Health Care*. 2017;46(3):215-29.
- [49] Drotar D, Ievers C. Age differences in parent and child responsibilities for management of cystic fibrosis and insulin-dependent diabetes mellitus. *J Dev Behav Pediatr*. 1994;15(4):265-72.
- [50] Friedemann-Sánchez G, Capistrant BD, Ron J, Novak L, Zuijdwijk C, Ogle GD, et al. Caregiving for children with type 1 diabetes and clinical outcomes in central India: The IDREAM study. *Pediatr Diabetes*. 2018;19(3):527-33.
- [51] Geffken GR, Lehmkuhl H, Walker KN, Storch EA, Heidgerken AD, Lewin A, et al. Family functioning processes and diabetic ketoacidosis in youths with type I diabetes. *Rehabil Psychol*. 2008;53(2):231-7.
- [52] Gowers SG, Jones JC, Kiana S, North CD, Price DA. Family functioning: a correlate of diabetic control? *J Child Psychol Psychiatry*. 1995;36(6):993-1001.
- [53] Grabill KM, Geffken GR, Duke D, Lewin A, Williams L, Storch E, et al. Family functioning and adherence in youth with type 1 diabetes: A latent growth model of glycemic control. *Child Health Care*. 2010;39(4):279-95.
- [54] Grey M, Jaser SS, Whittlemore R, Jeon S, Lindemann E. Coping skills training for parents of children with type 1 diabetes: 12-month outcomes. *Nurs Res*. 2011;60(3):173-81.
- [55] Hanna KM, Dashiff CJ, Stump TE, Weaver MT. Parent-adolescent dyads: association of parental autonomy support and parent-adolescent shared diabetes care responsibility. *Child Care Health Dev*. 2013;39(5):695-702.
- [56] Hanna KM, Dimeglio LA, Fortenberry JD. Parent and adolescent versions of the diabetes-specific parental support for adolescents' autonomy scale: development and initial testing. *J Pediatr Psychol*. 2005;30(3):257-71.
- [57] Hanna KM, DiMeglio LA, Fortenberry JD. Brief Report: Initial Testing of Scales Measuring Parent and Adolescent Perceptions of Adolescents' Assumption of Diabetes Management. *J Pediatr Psychol*. 2007;32(3):245-9.
- [58] Hanna KM, Guthrie D. Adolescents' behavioral autonomy related to diabetes management and adolescent activities/rules. *Diabetes Educ*. 2003;29(2):283-91.
- [59] Hanna KM, Juarez B, Lenss SS, Guthrie D. Parent-adolescent communication and support for diabetes management as reported by adolescents with type 1 diabetes. *Issues Compr Pediatr Nurs*. 2003;26(3):145-58.
- [60] Holmes CS, Chen R, Streisand R, Marschall DE, Souter S, Swift EE, et al. Predictors of youth diabetes care behaviors and metabolic control: a structural equation modeling approach. *J Pediatr Psychol*. 2006;31(8):770-84.
- [61] Hsin O, La Greca AM, Valenzuela J, Moine CT, Delamater A. Adherence and glycemic control among Hispanic youth with type 1 diabetes: role of family involvement and acculturation. *J Pediatr Psychol*. 2010;35(2):156-66.
- [62] Ingerski LM, Anderson BJ, Dolan LM, Hood KK. Blood glucose monitoring and glycemic control in adolescence: contribution of diabetes-specific responsibility and family conflict. *J Adolesc Health*. 2010;47(2):191-7.
- [63] Katz ML, Volkening LK, Butler DA, Anderson BJ, Laffel LM. Family-based psychoeducation and Care Ambassador intervention to improve glycemic control in youth with type 1 diabetes: a randomized trial. *Pediatr Diabetes*. 2014;15(2):142-50.
- [64] Kaugars AS, Kichler JC, Alemzadeh R. Assessing readiness to change the balance of responsibility for managing type 1 diabetes mellitus: adolescent, mother, and father perspectives. *Pediatr Diabetes*. 2011;12(6):547-55.

- [65] Kichler JC, Kaugars AS, Ellis J, Alemzadeh R. Exploring self-management characteristics in youths with type 1 diabetes mellitus: does membership in a glycemic control category matter? *Pediatr Diabetes*. 2010;11(8):536-43.
- [66] Kichler JC, Kaugars AS, Marik P, Nabors L, Alemzadeh R. Effectiveness of groups for adolescents with type 1 diabetes mellitus and their parents. *Fam Syst Health*. 2013;31(3):280-93.
- [67] Kichler JC, Seid M, Crandell J, Maahs DM, Bishop FK, Driscoll KA, et al. The Flexible Lifestyle Empowering Change (FLEX) intervention for self-management in adolescents with type 1 diabetes: Trial design and baseline characteristics. *Contemp Clin Trials*. 2018;66:64-73.
- [68] King PS, Berg CA, Butner J, Butler JM, Wiebe DJ. Longitudinal trajectories of parental involvement in Type 1 diabetes and adolescents' adherence. *Health Psychol*. 2014;33(5):424-32.
- [69] King PS, Berg CA, Butner J, Drew LM, Foster C, Donaldson D, et al. Longitudinal trajectories of metabolic control across adolescence: associations with parental involvement, adolescents' psychosocial maturity, and health care utilization. *J Adolesc Health*. 2012;50(5):491-6.
- [70] Landers SE, Friedrich EA, Jawad AF, Miller VA. Examining the interaction of parental involvement and parenting style in predicting adherence in youth with type 1 diabetes. *Fam Syst Health*. 2016;34(1):41-50.
- [71] Lee SL, Lo FS, Lee YJ, Chen BH, Wang RH. Predictors of Glycemic Control in Adolescents of Various Age Groups With Type 1 Diabetes. *J Nurs Res*. 2015;23(4):271-9.
- [72] Lehmkuhl HD, Cammarata C, Meyer K, Duke D, Lewin A, Williams L, et al. Psychometric properties of a self-report measure of adherence to the diabetes regimen. *Child Health Care*. 2009;38(1):76-90.
- [73] Leonard BJ, Kratz BJ, Skay CL, Rheinberger MM. Comparison of mother-father perceptions of their child's self-management of diabetes. *Issues Compr Pediatr Nurs*. 1997;20(2):69-87.
- [74] Leonard BJ, Skay CL, Rheinberger MM. Self-management development in children and adolescents with diabetes: the role of maternal self-efficacy and conflict. *J Pediatr Nurs*. 1998;13(4):224-33.
- [75] Mansfield AK, Addis ME, Laffel LM, Anderson BJ. Gender differences in reports of self-reliance for diabetes tasks in a pediatric sample. *International Journal of Men's Health*. 2004;3(1).
- [76] Marker AM, Noser AE, Clements MA, Patton SR. Shared Responsibility for Type 1 Diabetes Care Is Associated With Glycemic Variability and Risk of Glycemic Excursions in Youth. *J Pediatr Psychol*. 2018;43(1):61-71.
- [77] Marrero DG, Vandagriff JL, Kronz K, Fineberg NS, Golden MP, Gray D, et al. Using telecommunication technology to manage children with diabetes: the Computer-Linked Outpatient Clinic (CLOC) Study. *Diabetes Educ*. 1995;21(4):313-9.
- [78] McNabb WL, Quinn MT, Murphy DM, Thorp FK, Cook S. Increasing children's responsibility for diabetes self-care: the In Control study. *Diabetes Educ*. 1994;20(2):121-4.
- [79] Moreland EC, Tovar A, Zuehlke JB, Butler DA, Milaszewski K, Laffel LM. The impact of physiological, therapeutic and psychosocial variables on glycemic control in youth with type 1 diabetes mellitus. *J Pediatr Endocrinol Metab*. 2004;17(11):1533-44.
- [80] Mulvaney SA, Lilley JS, Cavanaugh KL, Pittel EJ, Rothman RL. Validation of the diabetes numeracy test with adolescents with type 1 diabetes. *J Health Commun*. 2013;18(7):795-804.
- [81] Murphy HR, Wadham C, Hassler-Hurst J, Rayman G, Skinner TC. Randomized trial of a diabetes self-management education and family teamwork intervention in adolescents with Type 1 diabetes. *Diabet Med*. 2012;29(8):e249-54.
- [82] Murphy HR, Wadham C, Rayman G, Skinner TC. Approaches to integrating paediatric diabetes care and structured education: experiences from the Families, Adolescents, and Children's Teamwork Study (FACTS). *Diabet Med*. 2007;24(11):1261-8.
- [83] Naar-King S, Ellis DA, Idalski A, Frey MA, Cunningham P. Multisystemic therapy decreases parental overestimation of adolescent responsibility for type 1 diabetes management in urban youth. *Fam Syst Health*. 2007;25(2):178-89.
- [84] Nakamura N, Sasaki N, Kida K, Matsuura N. Health-related and diabetes-related quality of life in Japanese children and adolescents with type 1 and type 2 diabetes. *Pediatr Int*. 2010;52(2):224-9.
- [85] Nansel TR, Rovner AJ, Haynie D, Iannotti RJ, Simons-Morton B, Wysocki T, et al. Development and validation of the collaborative parent involvement scale for youths with type 1 diabetes. *J Pediatr Psychol*. 2009;34(1):30-40.
- [86] Neumark-Sztainer D, Patterson J, Mellin A, Ackard DM, Utter J, Story M, et al. Weight control practices and disordered eating behaviors among adolescent females and males with type 1 diabetes: associations with sociodemographics, weight concerns, familial factors, and metabolic outcomes. *Diabetes Care*. 2002;25(8):1289-96.
- [87] Ott J, Greening L, Palardy N, Holderby A, DeBell WK. Self-efficacy as a mediator variable for adolescents' adherence to treatment for insulin-dependent diabetes mellitus. *Child Health Care*. 2000;29(1):47-63.

- [88] Palmer DL, Berg CA, Butler J, Fortenberry K, Murray M, Lindsay R, et al. Mothers', fathers', and children's perceptions of parental diabetes responsibility in adolescence: examining the roles of age, pubertal status, and efficacy. *J Pediatr Psychol.* 2009;34(2):195-204.
- [89] Palmer DL, Berg CA, Wiebe DJ, Beveridge RM, Korbel CD, Upchurch R, et al. The role of autonomy and pubertal status in understanding age differences in maternal involvement in diabetes responsibility across adolescence. *J Pediatr Psychol.* 2004;29(1):35-46.
- [90] Reid GJ, Dubow EF, Carey TC, Dura JR. Contribution of coping to medical adjustment and treatment responsibility among children and adolescents with diabetes. *J Dev Behav Pediatr.* 1994.
- [91] Robinson EM, Iannotti RJ, Schneider S, Nansel TR, Haynie DL, Sobel DO. Parenting goals: predictors of parent involvement in disease management of children with type 1 diabetes. *J Child Health Care.* 2011;15(3):199-209.
- [92] Rohan JM, Huang B, Pendley JS, Delamater A, Dolan L, Reeves G, et al. Predicting Health Resilience in Pediatric Type 1 Diabetes: A Test of the Resilience Model Framework. *J Pediatr Psychol.* 2015;40(9):956-67.
- [93] Siminerio LM, Charron-Prochownik D, Banion C, Schreiner B. Comparing outpatient and inpatient diabetes education for newly diagnosed pediatric patients. *Diabetes Educ.* 1999;25(6):895-906.
- [94] Srinivasan S, Craig ME, Beeney L, Hayes R, Harkin N, Ambler GR, et al. An ambulatory stabilisation program for children with newly diagnosed type 1 diabetes. *Med J Aust.* 2004;180(6):277-80.
- [95] Streisand R, Swift E, Wickmark T, Chen R, Holmes CS. Pediatric parenting stress among parents of children with type 1 diabetes: the role of self-efficacy, responsibility, and fear. *J Pediatr Psychol.* 2005;30(6):513-21.
- [96] Telo GH, Volkening LK, Butler DA, Laffel LM. Salient Characteristics of Youth with Type 1 Diabetes Initiating Continuous Glucose Monitoring. *Diabetes Technol Ther.* 2015.
- [97] Turner SL, Berg CA, Butner JE, Wiebe DJ. Attention Problems as a Predictor of Type 1 Diabetes Adherence and Metabolic Control Across Adolescence. *J Pediatr Psychol.* 2018;43(1):72-82.
- [98] Vesco AT, Anderson BJ, Laffel LM, Dolan LM, Ingerski LM, Hood KK. Responsibility sharing between adolescents with type 1 diabetes and their caregivers: importance of adolescent perceptions on diabetes management and control. *J Pediatr Psychol.* 2010;35(10):1168-77.
- [99] Weissberg-Benchell J, Glasgow A. The role of temperament in children with insulin-dependent diabetes mellitus. *J Pediatr Psychol.* 1997;22(6):795-809.
- [100] Weissberg-Benchell J, Goodman SS, Antisdell Lomaglio J, Zebracki K. The use of Continuous Subcutaneous Insulin Infusion (CSII): parental and professional perceptions of self-care mastery and autonomy in children and adolescents. *J Pediatr Psychol.* 2007;32(10):1196-202.
- [101] Weissberg-Benchell J, Nansel T, Holmbeck G, Chen R, Anderson B, Wysocki T, et al. Generic and diabetes-specific parent-child behaviors and quality of life among youth with type 1 diabetes. *J Pediatr Psychol.* 2009;34(9):977-88.
- [102] Wiebe DJ, Berg CA, Korbel C, Palmer DL, Beveridge RM, Upchurch R, et al. Children's appraisals of maternal involvement in coping with diabetes: enhancing our understanding of adherence, metabolic control, and quality of life across adolescence. *J Pediatr Psychol.* 2005;30(2):167-78.
- [103] Wiebe DJ, Chow CM, Palmer DL, Butner J, Butler JM, Osborn P, et al. Developmental processes associated with longitudinal declines in parental responsibility and adherence to type 1 diabetes management across adolescence. *J Pediatr Psychol.* 2014;39(5):532-41.
- [104] Wiebe DJ, Croom A, Fortenberry KT, Butner J, Butler J, Swinyard MT, et al. Parental involvement buffers associations between pump duration and metabolic control among adolescents with type 1 diabetes. *J Pediatr Psychol.* 2010;35(10):1152-60.
- [105] Wilksch SM, Starkey K, Gannoni A, Kelly T, Wade TD. Interactive programme to enhance protective factors for eating disorders in girls with type 1 diabetes. *Early Interv Psychiatry.* 2013;7(3):315-21.
- [106] Wu YP, Hilliard ME, Rausch J, Dolan LM, Hood KK. Family involvement with the diabetes regimen in young people: the role of adolescent depressive symptoms. *Diabet Med.* 2013;30(5):596-602.
- [107] Wu YP, Rausch J, Rohan JM, Hood KK, Pendley JS, Delamater A, et al. Autonomy support and responsibility-sharing predict blood glucose monitoring frequency among youth with diabetes. *Health Psychol.* 2014;33(10):1224-31.
- [108] Wysocki T, Nansel TR, Holmbeck GN, Chen R, Laffel L, Anderson BJ, et al. Collaborative involvement of primary and secondary caregivers: associations with youths' diabetes outcomes. *J Pediatr Psychol.* 2009;34(8):869-81.
- [109] Wysocki T, Taylor A, Hough BS, Linscheid TR, Yeates KO, Naglieri JA. Deviation from developmentally appropriate self-care autonomy. Association with diabetes outcomes. *Diabetes Care.* 1996;19(2):119-25.
- [110] Rubin R, Young-Hyman D, Peyrot M. Parent-child responsibility and conflict in diabetes care. *Diabetes.* 1989;38:28.

- [111] Ruggiero L, Mindell J, Kairys S. A new measure for assessing diabetes self-management responsibility in pediatric populations: The Diabetes Regimen Responsibility Scale. *The Behavior Therapist*. 1991;14:233-4.
- [112] Saletsky RD. Autonomy and mother-child relations in a sample of adolescents with insulin-dependent diabetes 1992.
- [113] Wysocki T, Meinhold P, Cox DJ, Clarke WL. Survey of diabetes professionals regarding developmental changes in diabetes self-care. *Diabetes Care*. 1990;13(1):65-8.
- [114] Wysocki T, Meinhold PA, Abrams KC, Barnard MU, Clarke WL, Bellando BJ, et al. Parental and professional estimates of self-care independence of children and adolescents with IDDM. *Diabetes Care*. 1992;15(1):43-52.
- [115] Laffel LM, Connell A, Vangsness L, Goebel-Fabbri A, Mansfield A, Anderson BJ. General quality of life in youth with type 1 diabetes: relationship to patient management and diabetes-specific family conflict. *Diabetes Care*. 2003;26(11):3067-73.
- [116] Noorani M, Ramaiya K, Manji K. Glycaemic control in type 1 diabetes mellitus among children and adolescents in a resource limited setting in Dar es Salaam - Tanzania. *BMC Endocr Disord*. 2016;16(1):29.
- [117] Hood KK, Butler DA, Anderson BJ, Laffel LMB. Updated and revised Diabetes Family Conflict Scale. *Diabetes Care*. 2007;30(7):1764-9.
- [118] de Wit M, Winterdijk P, Aanstoot H-J, Anderson B, Danne T, Deeb L, et al. Assessing diabetes-related quality of life of youth with type 1 diabetes in routine clinical care: the MIND Youth Questionnaire (MY-Q). *Pediatr Diabetes*. 2012;13(8):638-46.
- [119] Palmer DL, Osborn P, King PS, Berg CA, Butler J, Butner J, et al. The structure of parental involvement and relations to disease management for youth with type 1 diabetes. *J Pediatr Psychol*. 2011;36(5):596-605.
- [120] Wiebe DJ, Berg CA, Fortenberry KT, Sirstins J, Lindsay R, Donaldson D, et al. Physician recommendations about maternal involvement in adolescent diabetes management. *Diabetes Care*. 2008;31(4):690-2.
- [121] Anderson B, Ho J, Brackett J, Finkelstein D, Laffel L. Parental involvement in diabetes management tasks: relationships to blood glucose monitoring adherence and metabolic control in young adolescents with insulin-dependent diabetes mellitus. *J Pediatr*. 1997;130(2):257-65.
- [122] Schmidt C. Self-care in children with type 1 diabetes: a survey of mothers. *MCN Am J Matern Child Nurs*. 2007;32(4):223-9.
- [123] Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, De Vet H, et al. COSMIN study design checklist for Patient-reported outcome measurement instruments: Amsterdam, The Netherlands; 2019.
- [124] U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4:79.

APPENDIX A

Criteria to evaluate content validity^a

Psychometric feature	Criteria	Information extracted from the original study ^a	Evaluation (overall rating of the original study and reviewers' rating regarding review construct of interest)
Content validity	<i>Relevance</i>		
	1. Are the included items relevant for the construct of interest	a. Description of the construct of interest b. Description of the origin of the construct c. How were items developed; information as to what extent items refer to the construct of interest according to concept elicitation, literature, or professionals	<i>Rating original study</i> a. Extent to which information regarding the construct of interest is provided and is clear b. Extent to which information regarding the origin of the construct is provided and is clear c. Extent of evidence that is provided that items refer to the construct of interest according to concept elicitation, literature or professionals <i>Reviewers' rating</i> Extent to which items are relevant for the construct of interest
	2. Are the included items relevant for the target population of interest	a. Description of the target population b. Information whether the target population was represented in the elicitation of relevant items c. Methodological information regarding concept elicitation (i.e., type of data collection, background of moderators, use of topic or interview guide, information regarding data processing and coding process, sample size)	<i>Rating original study</i> a. Extent to which information regarding the target population is provided and is clear b. Extent to which the target population was represented in the elicitation of relevant items c. Extent to which the concept elicitation was adequate (i.e., appropriate qualitative data collection, skilled moderators, based on appropriate topic or interview guide, use of recording and transcripts, use of justified approach for analysis, extent to which data was coded by two independent researchers, data saturation, $n \geq 100$ for studies using surveys) <i>Reviewers' rating</i> Extent to which the items are relevant for the population of interest
	3. Are the included items relevant for the context of interest	Description of the context of interest	<i>Rating original study</i> Extent to which information is provided and is clear <i>Reviewers' rating</i> Extent to which the items are relevant for the context of interest
	4. Are the response options appropriate	Justification for the response options	<i>Rating original study</i> Extent to which the justification is provided and is clear

Psychometric feature	Criteria	Information extracted from the original study ^a	Evaluation (overall rating of the original study and reviewers' rating regarding review construct of interest)
			<i>Reviewers' rating</i> Extent to which the response options are appropriate for the construct, population, and context of interest
	5. Is the recall period appropriate	Justification for the recall period	<i>Rating original study</i> Extent to which the justification is available and is clear <i>Reviewers' rating</i> Extent to which the recall period is appropriate for the construct, population, and context of interest
	<i>Comprehensiveness</i>		
	6. Are no key concepts missing	a. Results of the comprehensiveness study among the population of interest b. Methodological information about the comprehensiveness study among the population of interest (i.e., instrument that was tested, type of data collection, sample size, background of moderators, use of topic or interview guide, information regarding data processing and coding process, adaptation based on results)	<i>Rating original study</i> a. Population of interest was asked about the comprehensiveness of the instrument; no key concepts were missing b. Extent to which the quality of the comprehensiveness study was adequate (i.e., final instrument tested, appropriate qualitative data collection, $n \geq 50$ for studies using surveys/ $n \geq 7$ in qualitative studies, skilled moderators, based on appropriate topic or interview guide, use of recording and transcripts, use of justified approach for analysis, two independent researchers involved, data saturation, were problems regarding comprehensiveness appropriately addressed by adapting the instrument) <i>Reviewers' rating</i> Extent of comprehensiveness for the construct, population and context of interest
	<i>Comprehensibility</i>		
	7. Are instructions understood by the population of interest as intended	a. Results of the comprehensibility study of the instructions according to the population of interest b. Methodological information about the comprehensibility study of the instructions among the population of interest (i.e., instrument that was tested, type of data collection, sample size,	<i>Rating original study</i> a. Population of interest was asked about the comprehensibility of the instructions; problems were adequately addressed b. Extent to which the quality of the comprehensibility study was adequate (i.e., final instrument tested, appropriate qualitative data collection, $n \geq 50$ for studies using surveys/ $n \geq 7$ in qualitative studies, skilled moderators, based on appropriate topic or interview guide, use of recording and transcripts, use of justified approach for analysis, two independent researchers involved, data saturation, were problems

Psychometric feature	Criteria	Information extracted from the original study ^a	Evaluation (overall rating of the original study and reviewers' rating regarding review construct of interest)
		background of moderators, use of topic or interview guide, information regarding data processing and coding process, adaptation based on results)	regarding comprehensibility appropriately addressed by adapting the instrument)
	8. Are items and response options understood by the population of interest as intended ^b	a. Results of the comprehensibility study of the items and response options among the population of interest b. Methodological information about the comprehensibility study of the items and response options among the population of interest (i.e., instrument that was tested, type of data collection, sample size, background of moderators, use of topic or interview guide, information regarding data processing and coding process, adaptation based on results)	<i>Rating original study</i> a. Population of interest was asked about the comprehensibility of the items and response options; problems were adequately addressed b. Extent to which the quality of the comprehensibility study was adequate (i.e., final instrument tested, appropriate qualitative data collection, $n \geq 50$ for studies using surveys/ $n \geq 7$ in qualitative studies, skilled moderators, based on appropriate topic or interview guide, use of recording and transcripts, use of justified approach for analysis, two independent researchers involved, data saturation, were problems regarding comprehensibility appropriately addressed by adapting the instrument)
	9. Are the items appropriately worded ^c		<i>Reviewers' rating</i> <ul style="list-style-type: none"> Extent to which the items and response options are appropriately worded (considering aspects such as reading level, ambiguous items, double barreled questions, jargon, value laden words, and length of items)
	10. Do the response options match the question ^c		<i>Reviewers' rating</i> <ul style="list-style-type: none"> Extent to which the response options match the question

^a As no content validity studies were identified among the included studies (i.e., study on the relevance, comprehensiveness and comprehensibility of an existing scale), these specific standards were not included in this table.

^b Questions 7 and 8 were based on the original study as the evaluation of these aspects cannot be rated by reviewers.

^c Questions 9 and 10 were only based on reviewers' ratings.

APPENDIX B

Criteria to evaluate the internal structure^a

Psychometric feature	Criteria	Information extracted from the included studies	Quality assessment
Internal structure	<i>Structural validity</i>		
	1. Criteria for factor structure <ul style="list-style-type: none"> • CFA: CFI or TLI or comparable measure >0.95 or RMSEA <0.06 or SRMR <0.08 • EFA: Not applicable 	<p>a. Methodological information (is the model based on a reflective model^b; rationale for conducting a factor analysis; type of analysis (CFA vs. EFA); n per item and sample size; EFA: data inspection technique (Barlett's tests and/or KMO measure); EFA: factor analytic method (PCA, PAF, ML); EFA: factor retention method; EFA: factor rotation method; EFA: used factor loading cutoff</p> <p>b. Results of the factor analysis (i.e., number of factors and number of items per factor; CFA: fit statistics like CFI, TLI, RMSEA; EFA: data inspection, i.e., Barlett's tests and/or KMO measure)</p>	<p>a. Standards (i.e., stated that the instrument is based on a reflective model^b; clear from rationale if unidimensionality^c or structural validity^d is of concern; CFA was performed; CFA: 1-to-7 participant-to-variable ratio and $n \geq 100$, EFA: 1-to-5 participant-to-variable ratio and $n > 200$; EFA: significant Barlett's test of Sphericity and/or > 0.60 KMO Measure of Sampling Adequacy; EFA: avoiding principal components analysis^e; EFA: factor analytic method with the fewest statistical assumptions while still providing all desired information; EFA: avoiding use of Kaiser criterion (Eigenvalue < 1), performance of a visual scree plot analysis and parallel analysis, Velicer's MAP (average squared correlation of the original data set) or both; EFA: rotation method that aligns with the expected correlation of the factors^f; EFA: item load > 0.40 primary factor, item load < 0.30 on alternative factors, item factor load difference > 0.20 between the primary factor and alternative factor loadings)</p> <p>b. Quality of evidence (i.e., extent of unexplained inconsistencies across studies, extent to which the included studies represent the population of interest)</p>
	<i>Internal consistency</i>		
2. Evidence for sufficient structural validity and Cronbach's alpha ≥ 0.70 for each unidimensional scale or subscale	<p>a. Methodological information (i.e., level of the analysis; type of analysis; methodological issues)</p> <p>b. Results of the internal consistency analysis per responder (i.e., reported Cronbach's alpha value; range of Cronbach's alpha across studies per responder; studies reporting Cronbach's alpha < 0.70)</p>	<p>a. Standards (i.e., internal consistency statistic calculated for each unidimensional scale or subscale; Cronbach's alpha was calculated; no methodological flaws)</p> <p>b. Quality of evidence (i.e., extent of unexplained inconsistencies across studies; total $n > 100$; extent to which the included studies represent the population of interest)</p>	

Abbreviations: EFA= Explanatory factor analysis; CFA= Confirmatory factor analysis; CFI= Comparative fit index; KMO= Kaiser-Meyer-Olkin; MAP= Minimum average partial; ML= Maximum likelihood; PAF= Principal axis factoring; PCA= Principal components analysis; RMSEA= Root mean square error of approximation; SRMR= Standard root mean residuals; TLI= Tucker-Lewis index.

^a As no item response theory or Rasch analyses were identified among the included studies, these specific standards were not included in this table.

^b Assumes that all items in a scale or subscale are manifestations of one underlying construct and are expected to be correlated).

^c Factor analysis is conducted to evaluate if the scale is unidimensional.

^d Factor analysis is conducted to evaluate the (expected) number of subscales in the instrument and the clustering of items within subscales.

^e Not a true form of factor analysis – does not account for the structure of correlations and does not make a distinction between common variance and unique variance.

^f No relation among factors: orthogonal rotation (i.e., quarimax and varimax); factors correlated: oblique rotation (i.e., promax and direct oblimin).

APPENDIX C

Criteria to evaluate the remaining measurement properties^a

Psychometric feature	Aspect	Criteria	Information extracted from the included studies	Quality of results
Remaining measurement properties	1. Reliability (inter-rater reliability)	ICC or Weighted Kappa ≥ 0.70	<ul style="list-style-type: none"> a. Methodological information (i.e., type of analysis; methodological issues) b. Results of the reliability analysis (i.e., reported Kappa of individual studies) 	<ul style="list-style-type: none"> a. Standards (i.e., for interviews inter-rater reliability analysis was performed; no methodological flaws) b. Quality of evidence (i.e., extent of unexplained inconsistencies across studies, total $n > 100$, extent to which the included studies represent the population of interest)
	2. Hypothesis test for construct validity	<ul style="list-style-type: none"> a. Results are in line in terms of direction (positive and negative) and magnitude (absolute or relative) with hypotheses regarding internal relationships and relationships with other instruments b. Results are in line in terms of direction (positive and negative) and magnitude (absolute or relative) with hypotheses regarding differences between relevant groups 	<ul style="list-style-type: none"> a. Comparison with another measurement instrument (convergent or concurrent validity) <ul style="list-style-type: none"> i. Methodological information (i.e., information regarding the comparator instrument; measurement properties of the comparator instrument; type of analysis) ii. Results of the comparison b. Comparison between subgroups (discriminative or known-groups validity) <ul style="list-style-type: none"> i. Methodological information (i.e., description of the characteristics of the subgroup; information regarding design and statistical methods) ii. Predefined hypothesis, results of the comparison 	<ul style="list-style-type: none"> a. Comparison with another measurement instrument (convergent or concurrent validity) <ul style="list-style-type: none"> i. Standards (i.e., construct measured by the comparator instrument is clear; sufficient measurement properties of the comparator instrument in a population similar to the population of interest; appropriate statistical methods) ii. Quality of evidence (i.e., extent of unexplained inconsistencies across studies, total $n > 100$, extent to which the included studies represent the population of interest, other measurements measure different constructs) b. Comparison between subgroups (discriminative or known-groups validity) <ul style="list-style-type: none"> i. Standards (adequate description of key characteristics of the subgroups; appropriate statistical methods) ii. Quality of evidence (i.e., extent of unexplained inconsistencies across studies, total $n > 100$, extent to which the included studies represent the population of interest, extreme groups were compared)

Abbreviations: ICC= Intraclass Correlation Coefficient.

^a As no test-retest reliability analyses were identified among the included studies, these specific standards were not included in this table.