

# **AACPDM Methodology to Develop Systematic Reviews of Treatment Interventions (Revision 1.2)**

**2008 Version**

---

**Inside**

**The Step-by-Step Process for Authors Creating an  
AACPD M Systematic Review**

**Approved by the Treatment Outcomes Committee  
December 17, 2008**

*Charlene Butler authored the first version of this manual in 1998-1999 and the 1998-1999 AACPD Treatment Outcomes Committee approved it. The purpose was to standardize the development and presentation of evidence tables and review articles about treatment outcomes for children and youth with developmental conditions.*

*The first revision (version 1.1) was approved by the Treatment Outcomes Committee at the September 2004 annual meeting. This document replaces the 2004 revision. The major change included in the 2008 revision is the addition of Levels of Evidence and conduct ratings for single subject design studies. Other changes were made to simplify the review process and make the review methodology easier to understand for reviewers.*

*Johanna Darrah PhD  
Robbin Hickman PT, DSc, PCS  
Maureen O'Donnell MD MSc  
Laura Vogtle PhD  
Lesley Wiart PhD*

## **INTRODUCTION**

The American Academy for Cerebral Palsy and Developmental Medicine (AAPDM) Treatment Outcomes Committee (TOC) developed a systematic review process to summarize the literature about specific intervention strategies used with children with developmental disabilities. Systematic reviews do not specify how to treat a condition; *rather* they gather and present the best evidence – for or against - the effectiveness of an intervention. Such reviews can also reveal a lack of evidence to make a judgment regarding treatment effectiveness. These reviews are not “best practice” documents or practice guidelines. The goal of these reviews is to present the evidence about interventions in an organized fashion. Such reviews can assist to identify gaps in evidence, and can help identify new research areas. The Academy is neither endorsing nor disapproving of an intervention in these reviews.

This document describes the updated methodology for developing and presenting an AAPDM systematic review – a review that is based on a two-part conceptual framework. This framework:

- 1) analyzes and categorizes treatment outcomes from studies according to the components of the International Classification of Functioning, Disability and Health (ICF) (WHO, 2001), and
- 2) judges the strength of the evidence from each article according to the study design and the researchers’ rigor in the conduct of the study.

Readers are encouraged to visit the website of the World Health Organization (WHO) to gain more information about the structure and intent of ICF, and to understand the terminology used in this systematic review methodology.

The methodology described here provides a road map for the essential steps in completing a systematic review using the AAPDM guidelines developed and endorsed by the TOC.

The following process involved in submitting a review to the TOC is covered in detail in this document:

- 1) Contact the Chair of the TOC to discuss and register your topic.
- 2) Follow the methodology for developing systematic reviews.
- 3) Submit your review to the TOC Chair for committee review.

Once the review is approved by the TOC, it will be posted on the AAPDM website. The review must also be signed off by the AAPDM president before submission to Developmental Medicine and Child Neurology for expedited review.

## **DEVELOPING AN AACPDM SYSTEMATIC REVIEW: STEP BY STEP**

### **1. Before you begin.**

Conducting a systematic review requires at least two authors/reviewers. When you have an idea for a systematic review, contact the TOC Chair. S/he will document your name and that of your co-author/s. The TOC Chair records the title and topic of your review to ensure that no similar review is in progress. At this point the Chair may provide some feedback regarding your review topic and whether it needs to be made more specific.

You can then proceed to complete the review. If you would like assistance with the methodology described here, TOC members would be happy to help you. You can contact committee members through the TOC Chair. It is essential that you contact the TOC Chair before you begin a review rather than after you have completed a review. The current TOC Chair can be contacted through the TOC section of the AACPDM website.

### **2. Define the population of interest.**

Define the patient population as precisely as possible to focus the literature search to relevant studies. See Example 1 below.

### **3. Define the intervention as specifically as possible.**

Specifically state the intervention in order to focus the literature search to relevant studies. See Example 1 below.

#### **Example 1: *Specifying the population and intervention addressed by a review. Populations specifically excluded must also be described.***

The intervention, *intrathecal baclofen*, includes baclofen administered by (1) single or multiple injections over the dorsal surface of the spinal cord or (2) a subcutaneous implanted pump that delivers a continuous infusion into the lumbar cerebrospinal fluid. It excludes orally-administered baclofen.

This review is concerned with children and youth (birth to 19 years) with cerebral palsy. *Cerebral palsy* has traditionally been described as an evolving disorder of motor function secondary to a non-progressive pathology of the immature brain and is characterized by abnormalities of movement (i.e., spasticity, athetosis, chorea, dystonia, and ataxia). Two-thirds of individuals with cerebral palsy have spasticity, either alone or in combination with the other movement abnormalities. The review includes studies whose subjects were primarily individuals diagnosed with cerebral palsy with spasticity, alone or in combination with other types of abnormal movement. This review excludes studies of spasticity of spinal origin (e.g., multiple sclerosis or spinal cord injury) or of cerebral origin due primarily to other causes such as traumatic brain injury.

Defining the population and intervention carefully is a crucial step as this will affect the generalizability of the review.

#### 4. Create, execute and record search strategy

The comprehensive search for evidence is a key factor in a good systematic review. A comprehensive search of published literature **must** be undertaken. All literature published in English regarding the intervention and its application to the population of interest must be identified.

With today's comprehensive, complex electronic search mechanisms, it is crucial that a search strategy be designed and recorded as part of this review process. It is highly recommended that the search strategy be developed with the assistance of a health science librarian.

For these reviews, the literature search strategy is limited to published literature. Some systematic reviews include "grey literature" which is unpublished, non-journal sources such as results from dissertations or abstracts from scientific meetings that have not yet been published in paper format. While these can provide useful information, they are often difficult to find and may not have been subjected to the same level of peer review as published literature. These unpublished, non-peer reviewed sources are **NOT** included in the AACPDM reviews. Only original, peer reviewed literature published in scientific journals is included.

When searching the published literature, identify key words or Medical Subject Headings (MeSH) that reflect your intervention and population already defined above. State the inclusion and exclusion criteria applied to the search and to the subsequent selection process.

Your search strategy, once developed, should be applied to electronic data bases which, given the nature of this field, should include Scopus, PubMed, MEDLINE and CINAHL. If the intervention is of an educational nature or educationally-related outcomes are possible, ERIC should also be searched. Similarly, PsycInfo should be searched if psychological interventions or outcomes are of interest. EMBASE should be searched if it is available to authors. The databases suggested are examples and do not represent an exhaustive list; authors are responsible for identifying appropriate databases to include in their search strategy. Dates used for the search strategy must also be included.

Authors could also review the following or similar sources to ensure that all relevant literature is included:

- Database of Reviews of Effectiveness (DARE): DARE summarizes reviews identified and appraised by National Health Service's Centre for Reviews and Dissemination at York in England.
- Cochrane Database of Systematic Reviews contains full text of systematic reviews completed by the Cochrane Collaboration, an international network committed to preparing and disseminating systematic reviews in health care
- The Physiotherapy Evidence Database (PEDro)
- OT Seeker

It is important to keep a careful record of the search strategy and to report it in the systematic review when it is submitted to the TOC for review. In addition, authors must

review the references of all retrieved articles for additional relevant literature. The same search strategy may be used in the future to update your systematic review. An example of a search strategy from a published TOC systematic review is provided below (example 2).

**Example 2: *Specifying the search strategy used to identify articles for the review.***

**LITERATURE SEARCH**

The literature search included PubMed (1950 through April 2007), CINAHL (1982-April 2007), and Cochrane Database of Systematic Reviews for studies published in English. The electronic search terms were (osteogenesis imperfecta AND (phosphonate OR bisphosphonate OR pamidronate OR alendronate OR residronate OR clodronate or etidronate OR olpadronate OR APD OR zoledronic acid). Reference lists in studies and review articles and researchers knowledgeable about this intervention were also consulted to identify potentially relevant studies. Abstracts and, if needed, full text of articles were reviewed to exclude publications which were not reports of treatment. Of 109 citations, 70 met inclusion criteria.

For guidance with searching for literature, we recommend that you review one of the many current evidence-based practice textbooks. Authors also can contact the Treatment Outcomes Committee members for assistance and tips. We strongly recommend that you solicit the assistance of a medical or health sciences librarian.

While one author of the review may complete the search process, all authors of the review must feel comfortable that the search was comprehensive. Therefore, all authors of the review should review the search strategy and search terms. If the number of titles identified in a search is not manageable (e.g., 500 or more), the search strategy should be reviewed with a librarian and the search terms modified to make the search more specific. Do not remove articles from the review based solely on their titles. It is better to refine the search terms if too many articles are identified. For most reviews, the identification of too many articles will not be a problem! Members of the TOC can help you with this process so please contact the TOC chair if you would like assistance.

Retrieve all abstracts identified in the search process. At least two authors must independently read the abstracts and independently decide if an abstract should be included in the review. If there is no abstract or if it is difficult to determine from the abstract if the paper meets the review criteria, retrieve and review the full paper. After reviewing the abstracts, all authors of the review meet to discuss their decisions about including articles and reach consensus. The process must be carefully documented and reported in the systematic review. Keep records of which studies were retrieved and which studies were excluded (along with the reasons for exclusion).

This review process includes studies that report empirical data, both group studies and single subject design studies. Even though qualitative studies may provide important information, they cannot currently be rated by our methodology. If qualitative studies are identified in the search, include them in the reference list and state that they were excluded from the review because they were qualitative studies. Readers of the review will then be able to obtain the citations to these articles from the reference list if they wish.

## 5. Extract data from each included study

At least two authors must participate in extraction of the data from the original papers that are to be included in the review. If more than two authors are involved in the review, the authors can work in pairs to review the literature. Summarize the extracted data on the “Study Data Extraction Summary Form”. Separate data summary forms are available for group studies and single subject design studies (**Appendix 1**). The steps in this data extraction process for EACH author are:

1. Independently read the identified articles.
2. Independently code the level of evidence based on the research design of the article. (described in **Appendix 2**, “Levels of Evidence”)
3. Independently extract subject and descriptive information relating to the definition of the population (described in **Appendix 3** “Constructing Table 2”)
4. Independently document the specific nature of the intervention of interest, as well as noting any description of co-interventions that may have occurred. (described in **Appendix 3** “Constructing Table 2”)
5. Independently assess the quality of the study and assign a quality rating. (described in **Appendix 4**, “Constructing Table 3”)
6. Independently identify outcomes of interest and the measures used to assess them (described in **Appendix 5** “Constructing Table 4”)
7. Independently code the ICF component represented by each outcome of interest (described in **Appendix 5** “Constructing Table 4”)
8. Independently identify any reported adverse events, number of cases reported for each adverse event and the methods used to identify them. (described in **Appendix 6** “Constructing Table 5”)
9. Authors will then come together, compare their data and reach consensus.

Retain the final summary sheets for each article in a secure location. It is important that all information is available for audit purposes in the event that a reader seeks clarification of information included in the review.

## **6. Use extracted data to create summary “evidence” tables.**

For each report, five tables will be completed and included in the paper. Only four tables will be included if adverse events have not been reported in the studies reviewed. Table 1 is a standard table defining the levels of evidence for group design (Table 1a) and single subject design studies (Table 1b). Table 2 summarizes information from all studies cited in the review, regardless of the level of evidence assigned to the study. Tables 3a (group studies) and 3b (single subject studies) provide the conduct rating for all studies with levels of evidence I, II or III. Studies assigned levels of evidence of IV or V are not included in this table. Table 4 is the ‘evidence table’ and summarizes the outcomes measured in each study by the ICF component represented and documents the results of the analyses. Only studies with levels of evidence I, II or III are included in this table. Table 5 is the adverse events table. Please consult the appropriate Appendix for further instructions on designing these tables. It is important to follow the standard format. It is very helpful to construct these tables before starting the discussion section.

## **7. Write the review article. The format of the article will be:**

Abstract

Introduction/Background

Operational Definition of Population and Intervention

Search Strategy

Organization of Evidence

Discussion

Conclusion/Future Directions



(i) Abstract

Follow the instructions of the journal. In most instances, you will submit the review to *Developmental Medicine and Child Neurology* and thus will follow their abstract guidelines.

(ii) Introduction

The introduction describes the intervention reviewed. It should be brief. It should also include the following 'boiler plate' description of the AACPDm process:

*The American Academy for Cerebral Palsy and Developmental Medicine (AACPDm) has undertaken the development of systematic reviews to summarize the literature about specific intervention strategies used to assist children with developmental disabilities. These reviews are not best-practice documents or practice guidelines, but rather they gather and present the best evidence – for and against – the effectiveness of an intervention. Their goal is to present the evidence about interventions in an organized fashion to identify gaps in evidence and help identify new research areas. The Academy is neither endorsing nor disapproving an intervention in these reviews. Every effort has been made to assure that AACPDm systematic reviews are free from real or perceived bias. Details of the disclosure and consensus process for AACPDm outcomes reports can be viewed at [www.AACPDm.org](http://www.AACPDm.org). Nevertheless, the data in an AACPDm Systematic Review can be interpreted differently, depending on people's perspectives. Please consider the conclusions presented carefully.*

(iii) Operational Definition of Population and Intervention

It is important that the population and the intervention be clearly described and defined.

(iv) Search Strategy

Describe the search strategy, including search terms and inclusion and exclusion criteria. The number of articles initially identified and the decision making process for removing articles needs to be explained. Remember to cite articles that were removed from the review because the researchers used qualitative methods.

(v) Organization of Evidence

Briefly describe the process of extracting the information and explain the content of each table so that the reader has an organizational 'map' of the review. It is important that the readers understand that only studies with levels of evidence of I, II or III are included in the final evidence table (Table 4). Each study should **not** be described in detail in the text - this information is already available in the summary table (Table 2).

(vi) Discussion

The discussion represents a synthesis of the information obtained from the articles. It is not the place to review each article individually but rather to discuss similarities and differences among the selected articles. The discussion section of each systematic review should follow the same process and answer 6 standard questions. When answering these questions, refer to the information available from studies that are included in the Evidence Table (Table 4). Information gleaned from studies assigned evidence levels IV or V is only discussed in questions 5 and 6, and not used to answer questions 1 to 4. It is vital that the discussion points are based on information from the more rigorous studies from both group and single subject design studies. These discussion questions represent the opportunity to discuss the results from the group and single subject design methods in an integrated fashion.

#### Question 1

##### **What evidence exists about the effects of the intervention on outcomes representing the ICF component of body functions and structures?**

Describe the outcomes of interest in Table 4 that represent the ICF component of body functions and structures. This is an opportunity to synthesize the results of the review. For example, authors could discuss if outcomes of interest at body functions and structures were similar or different across the studies. If the same outcomes were evaluated across studies, were the results similar or different? Is the same outcome of interest measured in different ways?

#### Question 2

##### **What evidence exists about the effects of the intervention in the ICF component of activity?**

Use a similar strategy as described above. If there are no outcomes of interest representing this ICF component, it is important to state this.

#### Question 3

##### **What evidence exists about the effects of the intervention in the ICF component of participation?**

See strategy for question 2.

#### Question 4

##### **What evidence exists for linkages of effects within and between these components?**

This is an opportunity to examine the linkages across the different components. Did any study systematically examine the relationships of outcomes representing different ICF components? Did any of the studies measure outcomes representing the component of body functions and structures assume that changes in outcomes representing this component would automatically improve outcomes representing activity and participation? This question represents an excellent opportunity to evaluate the assumptions that clinicians and researchers may be making about the relationships among outcomes representing different components of the ICF.

#### Question 5

##### **What kinds and magnitude of adverse events have been documented?**

An understanding of the medical risks is necessary for assessing whether the benefits of an intervention outweigh its risks. The review will report the type and severity of adverse events that have been reported. It will not express a judgment about whether an intervention does more good than harm because a benefit/harm analysis can more appropriately be made by the reader who is considering the intervention for a particular individual. Were there adverse events reported that require further evaluation? Were there credible systems for recording adverse events? Discussion regarding level IV and V studies must be identified as information gleaned from studies with less rigorous design methods. If no adverse events were mentioned, insert a statement under this question indicating that no adverse events were discussed in the articles. This section should include a statement that indicates that an absence of reported adverse events does not necessarily mean that none exist; studies may not be adequately powered to detect all adverse events.

#### Question 6

##### **What is the strength of the evidence?**

In this section, discuss the two types of articles (group research and single subject research) separately because the strength of group studies and single subject studies are evaluated

using different criteria. It is important to remember that this question still looks at the evidence as a whole, not at studies individually, and that the studies included in Table 4 (i.e. levels I, II and III) are discussed first when answering this question. At the end of this question, Level IV and V studies that provide interesting results or ideas for future research studies can be discussed, but it must be identified as information gleaned from studies with less rigorous design methods.

#### *Group Studies*

For the literature representing group design studies, the answer to this question can include an estimate of how many subjects in total are represented across all studies. Design issues such as sample size, power, and specificity and ability to replicate the intervention can be discussed. Comments can also include the clinical significance of the outcomes evaluated. Any concerns about omissions in the literature can also be discussed (e.g., no outcomes representing the ICF component of activity/participation).

#### *Single Subject Design Studies*

For the literature representing single subject design studies, the response to this question can include a discussion of the stability of the data in both intervention and baseline phases, e.g., considering changes in data trend, level (magnitude) and variability. Treatment fidelity and participant selection can also be discussed (i.e., was the treatment delivered appropriately and were the children treated appropriate for the intervention?).

#### *(vii) Conclusion/Future Directions*

This is an opportunity for the authors to summarize the important points and suggest future directions for research. It should be brief and not duplicate the discussion section. If findings from studies not included in the evidence table (Table 4) are included as areas for future consideration, it should be clearly stated that this suggestion originates from studies with less rigorous design methods.

### **8. Submit the article to the Treatment Outcomes Committee for publication under the imprimatur of the AACPDM.**

In order for a review using this methodology to be published in *Developmental Medicine and Child Neurology* under the imprimatur of the AACPDM, the Treatment Outcomes Committee must review the article. Do not submit your review directly to DMCN without first sending it to the TOC Committee. When you have completed your review, submit your manuscript to the TOC chair. It will then be reviewed by five past or current members of the TOC including at least three active TOC members. You will receive feedback from this review group through the TOC Chair. The role of the committee is to serve in a reviewer capacity and to ensure the review has followed the methodology outlined here.

When the TOC review group has ‘signed off’ the review, and you, as authors, are happy with the review, the Chair will forward the review to the President of the AACPDM who reviews it prior to it being forwarded to the DMCN with the committee’s approval. The goal is that this process then leads to an expedited review process for the paper at DMCN. At this stage, the review will also be posted on the AACPDM website.

Throughout this process, the TOC members are there to help you. At any stage please feel free to contact the TOC chair who will assist you or direct you to others who can assist.

**Study Data Extraction Summary Form for Group Design Studies**

**Reviewer's Name:**

**Citation information:**

---

**1. Level of Evidence** (derived from levels of evidence, Appendix 2)

Research Design:

Level of Evidence (based on that design):

**2. Conduct Rating of the Study** (derived from answering conduct rating questions, Appendix 4)

**3. Descriptive Information about the Study**

Practice setting (place and/or type):

Participant description: (diagnosis, diagnostic subgroups, if relevant, ages severity, similarities between treatment and control groups, if relevant, etc.)

Number: In treatment group  
In control group

Specific intervention used: (nature, duration, dose, if relevant)

Description of control state (if used):

AACPDM SYSTEMATIC REVIEW OF THE EVIDENCE  
 APPENDIX 1: STUDY DATA EXTRACTION SUMMARY FORMS

**4. Outcomes**

<i>Outcome of Interest</i>	<i>Measure Used to Assess</i>	<i>ICF Component</i>	<i>Results</i>

Legend:

ICF Component      Component of ICF represented by measure and outcome of interest (record after completing Step 6, pg. 7). For information about classifying outcomes by the ICF component represented, please refer to Appendix 5 “Constructing Table 4”.

Results              Report results using the following three abbreviations:  
*ss: Statistically significant. Record the ‘p’ value. If there are two groups, record if the results favored the intervention or the control group*  
*ns: No statistical significance*  
*nr: Results of statistical tests not reported*

**5. Adverse Events**

Adverse events reported (number of cases)

Method of ascertaining adverse event (not stated; longitudinal registry in a defined sample or population; active, systematic surveillance)

**Study Data Extraction Summary Form for Single Subject Research Design Studies**

***Reviewer's Name:***

***Citation information:***

---

**1. Level of Evidence** (derived from levels of evidence, Appendix 2)

Research Design:

Level of Evidence (based on that design):

**2. Conduct Rating of the Study** (derived from answering conduct rating questions, Appendix 4)

**3. Descriptive Information about the Study**

Practice setting: (Place and/or type):

Number of participants

Participant description: (diagnosis, ages severity)

Description of intervention: (nature of intervention, number of phases, length of each phase, etc.)

**4 Outcomes**

<b><i>Outcome of Interest</i></b>	<b><i>Measure Used to Assess</i></b>	<b><i>ICF Component</i></b>	<b><i>Results</i></b>

Legend:

ICF Component      Component of ICF represented by measure and outcome of interest.  
 For information about classifying outcomes by the ICF component represented, please refer to Appendix 5 “Constructing Table 4”.

Results              Report results using the following three abbreviations:  
*ss: Statistically significant.*  
*ns: No statistical significance*  
*nr: Results of statistical tests not reported*

**5. Adverse Events**

Adverse events reported (number of cases)

Method of ascertaining adverse event (not stated; active, systematic surveillance)

## **CODING LEVELS OF EVIDENCE**

### **Group Research Designs**

There are numerous systems in the literature developed by organizations interested in reviewing the quality of evidence for grading the “level of evidence”. This grading is the most important step in determining the quality of the study.

The classification used for the AACPDm reviews is based on the work of Dr. David Sackett. First published in 1980, the grading system was for many years referred to as “Sackett’s levels of evidence and grades of recommendation”. The classification was republished with little change in 1993 but, more recently, has evolved further and changed under the auspices of the National Health Service Research and Development Centre for Evidence Based Medicine (CEBM) in Oxford, England. The current version developed by Sackett (then Director of CEBM) and his colleagues was posted on the CEBM web site on the Internet at [www.cebm.net](http://www.cebm.net) in 2001.

This hierarchy of “levels of evidence” is based on research design types. The following table, *which will be called “Table 1a – Levels of Evidence for Group Designs” in your review*, shows the classification of levels of evidence for group studies that the AACPDm uses in its reviews. In descending order, the designs are decreasingly able to demonstrate that the intervention—and not something else—was responsible for the observed outcome. Level I evidence is the most definitive for establishing causality, with greatest reduction in bias. Level IV can only hint at causality; Level V only suggests the possibility.

<b>Level</b>	<b><i>Intervention (Group) studies</i></b>
<b>I</b>	Systematic review of randomized controlled trials (RCTs) Large RCT (with narrow confidence intervals) (n >100)
<b>II</b>	Smaller RCT’s (with wider confidence intervals) (n<100) Systematic reviews of cohort studies “Outcomes research” (very large ecologic studies)
<b>III</b>	Cohort studies (must have concurrent control group) Systematic reviews of case control studies
<b>IV</b>	Case series Cohort study without concurrent control group (e.g. with historical control group) Case-control Study
<b>V</b>	Expert Opinion Case study or report Bench research Expert opinion based on theory or physiologic research Common sense/anecdotes



AACPDM SYSTEMATIC REVIEWS OF THE EVIDENCE  
APPENDIX 2: Constructing Table 1

Authors of an AACPDM review should consult a general clinical epidemiology textbook prior to undertaking their review to ensure they are classifying studies appropriately.

A word of caution and example of error in study classification: Case series studies (i.e., one group of patients measured for a given outcome or state, then provided with an intervention and measured again) can be erroneously classified as case control studies in which the cases acted as their own controls. A case-control study involves identifying a group of individuals *with* a given state/poor outcome (cases) and a group *without* the given state/good outcome (controls) and then looking back historically to identify whether or not both groups were equally exposed to the intervention of interest (the exposure). This is one example of a pitfall in assigning level of evidence, demonstrating the need to understand study design prior to undertaking the review process. In psychology and education, case series studies are defined as one-group, pretest-posttest designs.

### Single Subject Research Designs (SSRD)

This hierarchy of “levels of evidence” is based on research design types. The following table, **which will be called “Table 1b – Levels of Evidence for Single Subject Research Designs”** in your review, shows the classification of levels of evidence for single subject studies used by the AACPDm its reviews. Only insert this table if your review contains studies using single subject design methodology. These levels were developed by members of a subcommittee (Lynne Romeiser Logan, Robbin Hickman, Susan Harris, Carolyn Heriza) under the direction of Lynne Romeiser Logan, a TOC member. This work is published in *Developmental Medicine and Child Neurology*, 2008, 50: 99-103. For additional information, please refer to this article.

<b>Level</b>	<b>Single subject design studies</b>
<b>I</b>	Randomized controlled N-of-1 (RCT), alternating treatment design (ATD), and concurrent or non-concurrent multiple baseline design (MBDs); generalizability if the ATD is replicated across three or more subjects and the MBD consists of a minimum of three subjects, behaviors, or settings. These designs can provide causal inferences.
<b>II</b>	Non-randomized, controlled, concurrent MBD; generalizability if design consists of a minimum of three subjects, behaviors, or settings. Limited causal inferences.
<b>III</b>	Non-randomized, non-concurrent, controlled MBD; generalizability if design consists of a minimum of three subjects, behaviors or settings. Limited causal inferences.
<b>IV</b>	Non-randomized, controlled SSRDs with at least three phases (ABA, ABAB, BAB, etc.); generalizability if replicated across three or more different subjects. Only hints at causal inferences.
<b>V</b>	Non-randomized controlled AB SSRD; generalizability if replicated across three or more different subjects. Suggests causal inferences allowing for testing of ideas.

AACPDM SYSTEMATIC REVIEWS OF THE EVIDENCE  
APPENDIX 3: Constructing Table 2

Table 2 summarizes all the studies that are included in the review, regardless of the levels of evidence. Both group design studies and single subject design studies are included in this table but in different sections. For each section (group and single subject), studies are listed in chronological order. The first column contains the year of publication and the first author's surname with a superscript citation to the reference list. The second column contains the level of evidence and the research design type of each study (from Table 1). It also provides the conduct rating of the study for level I, II, and III studies (see appendix 4). The third column describes the parameters of the population represented by the sample; this information is derived from the inclusion/exclusion criteria described for participant selection in the individual articles. The fourth column provides the total number of subjects in each group or the total number of participants in the single subject design study. The ages of the participants are provided in column 5. Column 6 provides a description of the intervention evaluated and column 7 provides information about the control intervention (if there was one). For group studies, these columns contain information such as the location of the intervention and the frequency, intensity and duration of the intervention. For single subject design studies, there is only one intervention column which includes information regarding the length of the intervention, the number of phases, etc., as well as a description of the intervention.

Table 2 provides descriptive information for all studies in the review and consolidates information from various sources. This information does not have to be repeated in the text of the review. This is the only table that contains all studies in the review – it is an important table. It is strongly recommended that authors construct this table as they extract the information from the literature and before they begin the Discussion section.

The sample table on the next page provides examples of information for one group study and one single subject design study. The group study example is extracted from a review of conductive education. The single subject design study is a recent study in the literature but has not been included in an AACPD published review (Washington, Deitz, White & Schwartz, 2002. The effects of a contoured foam seat on postural alignment and upper-extremity function in infants with neuromotor impairments. *Physical Therapy*, 82: 1064-76).

**Table 2: Summary of studies – interventions and participants**

<b>Group Studies</b>	<b>Level of evidence and conduct rating</b>	<b>Participants</b>	<b>Total n</b>	<b>Ages</b>	<b>Intervention</b>	<b>Control Intervention</b>
<b>Research design</b>						
1972a Heal <sup>18</sup>	II-W (2/7) Cohort study with concurrent control group	CP, non-ambulatory, IQ<70, able to understand simple instructions Children in treatment and control groups matched on mental age, chronological age, type of CP and motor ability	25 CE=10 Ctl=15	5-13 y	CE by 2-4 professionally trained 'therapist-teachers' and institutional aides in a residential school; 13 1/2 hr/day Duration: 12 months	3 training programs in orthopedic residential schools Duration: 12 months
<b>Single Subject Studies</b>	<b>Level of evidence and conduct rating</b>	<b>Participants</b>	<b>Total n</b>	<b>Ages</b>	<b>Intervention</b>	
<b>Research design</b>						
2002 Washington	I-S (11/14) Alternating treatment time series. Treatments in intervention phase randomized	Spastic quadriplegia, hemiplegia, Down syndrome, congenital hypotonia. All receiving PT treatment, able to reach and grasp a toy in supported sitting, unable to sit independently as demonstrated by level of sitting value 3 or 4	4	9 to 18 months	A phase: Baseline data collected in standard high chair with no supports for 4 days over 2 separate 5-minute periods B/C phases: Intervention data collected for 2 separate 5-minute periods, one for each intervention (foam liner and CFS). Data collected daily for 8 days for each infant	

CP, cerebral palsy; CE, conductive education; PT, physical therapy; CFS, contoured foam seat

### **ASSESSING CONDUCT OF A STUDY**

Assessing the quality or conduct of an intervention study involves a number of steps. You have already completed the most important step by determining the “levels of evidence” classified by the study design. However, even if a study is deemed strong based on study design (level of evidence), it may still have problems that limit its quality. For this reason, researchers have developed “conduct” questions which can be used to rate study quality in addition to level of evidence.

A subgroup of the 2003/2004 committee (Maureen O'Donnell, Johanna Darragh, Richard Adams, Lori Roxborough and Diane Damiano) reviewed the literature regarding conduct/quality scoring systems and adapted those systems to create the following conduct scoring system for group design studies included in AACPDm reviews. **Please note that this quality assessment needs to be completed only for those studies with levels of evidence I, II or III. Studies with levels of evidence IV or V will not be in this table.** This table is entered in the manuscript as Table 3 (Table 3a if the review contains single subject design studies). Insert the conduct questions at the bottom of the table – the numbered columns correspond to the numbered questions. Each question should be answered “yes” (criterion/criteria present) or “no” (criterion/criteria not present). For group studies, the conduct of an individual study will be judged as Strong (‘yes’ score on 6-7 of the questions), Moderate (score 4 or 5) or Weak (score  $\leq 3$ ). Table 3b will contain the conduct rating for single subject design studies with levels of evidence I, II or III. This rating scale was designed by a group of experts in single subject design (Logan, Hickman, Harris and Heriza). For single subject design studies, the conduct of an individual study will be judged as Strong (‘yes’ score on 11-14 questions) Moderate (score 7-10) or Weak (score  $< 7$ ).

A sample of the table format follows as well as the conduct questions for both group design and single subject design studies.

**Table 3a Conduct of Group Design Studies**

<i>Study</i>	<i>Level/Quality</i>	1	2	3	4	5	6	7
1972a Heal <sup>18</sup>	II-W (2/7)	yes	no	no	no	n	yes	no
1995 Coleman <sup>16</sup>	II-W (4/7)	no	no	yes	yes	yes	yes	no
1995 Catanese <sup>14</sup>	II-W (4/7)	no	no	yes	yes	yes	yes	no
1995a Hur <sup>21</sup>	III-W (2/7)	no	yes	no	no	no	yes	no
1995b Hur <sup>22</sup>	III-W (2/7)	no	yes	no	no	no	yes	no
1997 Hur <sup>20</sup>	III-W (2/7)	yes	no	no	no	no	yes	no
1998 Reddihough <sup>24</sup>	I-S (6/7)	yes	no	yes	yes	yes	yes	yes

**Conduct Questions**

1. Were inclusion and exclusion criteria of the study population well described and followed?
2. Was the intervention well described and was there adherence to the intervention assignment? (for 2-group designs, was the control exposure also well described?) Both parts of the question need to be met to score 'yes'.
3. Were the measures used clearly described, valid and reliable for measuring the outcomes of interest?
4. Was the outcome assessor unaware of the intervention status of the participants (i.e., were the assessors masked)?
5. Did the authors conduct and report appropriate statistical evaluation including power calculations? Both parts of the question need to be met to score 'yes'.
6. Were dropout/loss to follow-up reported and less than 20%? For 2-group designs, was dropout balanced?
7. Considering the potential within the study design, were appropriate methods for controlling confounding variables and limiting potential biases used?

### **Table 3b Conduct Questions for Single Subject Design Studies**

The same format is followed for this table as in Table 2a. The conduct questions used to evaluate the quality of the studies are provided below:

1. Was/were the participant(s) sufficiently well described to allow comparison with other studies or with the reader's own patient population?
2. Were the independent variables operationally defined to allow replication?
3. Were intervention conditions operationally defined to allow replication?
4. Were the dependent variables operationally defined as dependent measures?
5. Was inter-rater or intra-rater reliability of the dependent measures assessed before and during each phase of the study?
6. Was the outcome assessor unaware of the phase of the study (intervention vs. control) in which the participant was involved?
7. Was stability of the data demonstrated in baseline, namely lack of variability or a trend opposite to the direction one would expect after application of the intervention?
8. Was the type of SSRD clearly and correctly stated, for example, A-B, multiple baseline across subjects?
9. Were there an adequate number of data points in each phase (minimum of five) for each participant?
10. Were the effects of the intervention replicated across three or more subjects?
11. Did the authors conduct and report appropriate visual analysis, for example, level, trend and variability?
12. Did the graphs used for visual analysis follow standard conventions, for example x- and y- axes labeled clearly and logically, phases clearly labeled (A,B, etc.) and delineated with vertical lines, data paths separated between phases, consistency of scales?
13. Did the authors report tests of statistical analysis, for example celeration line approach, two-standard deviation band method, C-statistic, or other?
14. Were all criteria met for the statistical analyses used?

AACPDM SYSTEMATIC REVIEWS OF THE EVIDENCE  
APPENDIX 4: Constructing Table 3

If your review of the literature includes a systematic review, evaluate the conduct of the review using the questions below. These questions have been taken from an article that reviewed the validity of this index (Oxman and Guyatt, 1991). The table should have the same format as Tables 3a and 3b and be labeled Table 3c. There is no grading of the total conduct score.

Questions to evaluate the conduct of a systematic review included in the review:

1. Were the search methods reported?
2. Was the search comprehensive?
3. Were the inclusion criteria reported?
4. Was selection bias avoided?
5. Were the validity criteria reported?
6. Was validity assessed properly?
7. Were the methods used to combine studies reported?
8. Were the findings combined appropriately?
9. Were the conclusions supported by the reported data?
10. What was the overall scientific quality of the overview?

If you require assistance in using this conduct rating, please refer to the reference below, or contact a member of the Treatment Outcomes Committee.

Oxman, A. D. and Guyatt, G. H. (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology*, **44**, 1271-1278.



AACPDM SYSTEMATIC REVIEW OF THE EVIDENCE  
APPENDIX 5: Constructing Table 4

**Table 4 Summary of studies: outcomes, measures, and results**

Table 4 is the ‘evidence’ table. It contains information regarding the outcomes measured in each study and the results of the analysis for each study. The AACPDm methodology categorizes outcomes using the components described in the International Classification of Functioning, Disability and Health (ICF), developed by the World Health Organization (WHO). This appendix contains a short description of the ICF components. We strongly recommend that any group conducting a review should access a copy of the textbook available from the WHO.

Only studies with a level of evidence of I, II or III are included in this table. As in Table 2, group design and single subject design studies are in separate sections of the table. In both sections, list studies in chronological order. **Code each outcome of interest by the ICF component represented. Then note the results of analysis in the appropriate ICF column. Report the statistical results as described on the summary extraction data sheet for each study type.**

The sample table provided shows one group study and one single subject design study. The contextual factors column may often be empty because few studies evaluate contextual factors as an outcome of interest. However, it is important to retain this column in the table, even if empty, because this is a strong reminder that contextual factors often are not evaluated. Outcomes that are appropriate for this column could be such things as ‘teacher attitudes’ or ‘physical accessibility.’

All abbreviations used in the table are noted at the bottom of the table in order of appearance in the table.

## **CODING OUTCOMES BY ICF COMPONENT**

### **BACKGROUND**

The AACPDm coding system uses the WHO International Classification of Functioning, Disability and Health (ICF), published September 2001. All systematic reviews must use this classification.

The ICF has *two parts*, each with *two components*. Each component can be described in positive or negative terms. Negative terminology is used if a deficit is present.

Part 1: Functioning and Disability

- (a) Body Functions and Structures (negative term = impairment)
- (b) Activities and Participation (negative terms = activity limitation and participation restriction)

Part 2: Contextual Factors

- (a) Environmental Factors (either facilitators or barriers)
- (b) Personal Factors

### **FUNCTIONING AND DISABILITY:**

Body Structures are anatomical parts of the body such as organs, limbs and their components. Impairments are problems in body structure as a result of deviation or loss. Structures include structures of the nervous system, the eye, structures involved in voice and speech, structures of the cardiovascular system, immunologic and respiratory systems, structures related to digestive, metabolic and endocrine systems, structures related to genitourinary and reproductive systems, structures related to movement, and skin-related structures.

Body Functions are the physiologic functions of body systems (including psychological functions). Impairments are problems in body function as a result of deviation or loss. This component includes mental function (e.g., intellect, temperament), sensory function (e.g., visual acuity, balance, pain), functions of the cardiovascular system, haematological, immunological and respiratory systems (e.g., heart rate, general physical endurance), function of the digestive, metabolic and endocrine systems (e.g. swallowing), function of the genitourinary and reproductive system (e.g., urinary incontinence) and neuro-musculoskeletal and movement-related functions (e.g., range of motion, muscle strength).

Activity is the execution of a task or action by an individual. Participation is involvement in a life situation. Activity limitations are difficulties an individual may have in executing activities. Participation restrictions are problems an individual may experience in involvement in life situations. WHO provides four different alternatives for coding activity and participation domains, ranging from using them interchangeably for all activities to

## AACPDM SYSTEMATIC REVIEW OF THE EVIDENCE

### APPENDIX 5: Constructing Table 4

having them represent mutually exclusive activities. For our purposes, we have chosen to have total overlap of domains for the two definitions. That is, when classifying outcomes of a study for Table 4, activity and participation outcomes will be identified in one heading “activities and participation”.

#### **CONTEXTUAL FACTORS**

Environmental factors make up the physical, social and attitudinal environment in which people live and conduct their lives. They can be viewed as facilitators (positive influence) or barriers (negative influence).

Personal factors are the particular background of an individual’s life and living and compromise features of the individual that are not part of a health condition. These factors may include gender, race, age and other health conditions. To date, these are not well developed in the ICF manual and must be used with caution, to ensure they are not better classified under body structure or body function.

#### **STICKY WICKETS IN CODING**

To develop the systematic review evidence tables, identify and code the measures found in the review of each research studies. Each measure will be coded as to whether it represents an outcome in the body functions/structures, activity/participation or contextual factors components.

If you are unclear as to how to classify a measure, look carefully at the items of the measure. For example, some believe that quality of life measures are at the level of participation; however, they too must be reviewed as the component items may in fact be at a body functions/structures or activity level. A scale or outcome measure should be reviewed for content and coded based on the majority of items. Subscales of a measure may represent different ICF components; however, these should be differently coded ONLY if the developers of the scale have demonstrated the reliability and validity of the sub-scales of the scale as stand-alone items.

#### **SUMMARY**

The ICF classification system represents a common language for rehabilitation disciplines to use to describe the level of client goals, intervention or outcome. For our purpose here, we are using it to code treatment outcome from studies of interventions in order to aggregate otherwise disparate research results into categories to help us make sense of bodies of evidence.

A complete description of these ICF components is found in the WHO publication from 2001 outlining the system. This is found and referenced at the WHO website found at <http://www.who.int/classifications/icf/en/>

AACPDM SYSTEMATIC REVIEWS OF EVIDENCE  
 APPENDIX 5: Constructing Table 4

Table 4: Summary of studies: outcomes, measures, and results

<b>Group Studies</b>	<b>Outcome of interest</b>	<b>Measure</b>	<b>Components of Health</b>		
			<b>Body Structure/s Body Functions</b>	<b>Activities and Participation</b>	<b>Contextual Factors</b>
Catanese <sup>14</sup> 1995	Gross motor	VAB-BR-Video		<i>p</i> <.01 CE	
		Video Ratings		<i>p</i> <.03 CE	
	Fine Motor	VAB-BR-Video		<i>ns</i>	
		Video Ratings		<i>p</i> <.01 CE	
	Receptive language	VAB-BR-Video		<i>ns</i>	
		Video Ratings		<i>ns</i>	
		VAB-CR		<i>ns</i>	
		VAB-BR-Video		<i>ns</i>	
	Expressive language	Video Ratings		<i>ns</i>	
		VAB-CR		<i>ns</i>	
	Grooming	VAB-BR-Video		<i>ns</i>	
		VAB-CR		<i>ns</i>	
	Feeding	VAB-BR-Video		<i>ns</i>	
		VAB-CR		<i>ns</i>	
	Dressing	VAB-CR		<i>ns</i>	
	Social Interaction	VAB-CR		<i>p</i> <.001 Ctl	
	Play	VAB-CR		<i>p</i> =.05 Ctl	
	Toileting	VAB-CR		<i>p</i> <.02 CE	
	Parent/Family Problems	QRS-F modified			<i>p</i> <.05 CE
	Pessimism	QRS-F modified			<i>ns</i>
Child Characteristics	QRS-F modified		<i>ns</i>		
ADL	Video Ratings		<i>p</i> =.01 CE		
Compliance	Video Ratings		<i>ns</i>		
Cognitive Ability	CMMS	<i>p</i> =.005 Ctl			
	RDLS	<i>ns</i>			
	PPVT	<i>ns</i>			
Cognitive and Physical Skills	WPPSI (drawing and maze subtests)	<i>p</i> <.03 Ctl			
Numeracy skills	SB subtest	<i>ns</i>			

AACPDM SYSTEMATIC REVIEWS OF EVIDENCE  
 APPENDIX 5: Constructing Table 4

<b>Single Subject Design Studies</b>	<b>Outcome of interest</b>	<b>Measure</b>	<b>Components of Health</b>		
			<b>Body Structure/s Body Functions</b>	<b>Activities and Participation</b>	<b>Contextual Factors</b>
Washington 2002	Postural Alignment	Frequency time intervals using videotapes of child in highchair with anatomical markers and visual guides	p=.008 demonstrating benefits for CFS for all 4 subjects		
	Engagement with toys	Percent of intervals with 2 hands on toy			<i>ns for all 4 subjects between 2 conditions</i>
		Percent of intervals with no hands on tray and 1 or 2 hands on toy			<i>p=.008 for one infant (+CFS)</i>

VAB-BR, Vulpe Assessment Battery-Behaviour rating; p, significance level; ns, not statistically significant; VAB-CR, Vulpe Assessment Battery-Caregiver Rating; QRS-F, Questionnaire on Resources and Stress (short form); CMMS, Columbia Mental Maturity Scale; SB, Stanford Binet; RDLS, Reynell Developmental Language Scale; PPVT, Peabody Picture Vocabulary Test; WPPSI, Wechsler Pre-school Scale of Intelligence; CFS, contoured foam seat.

AACPDM SYSTEMATIC REVIEWS OF EVIDENCE  
 APPENDIX 6: Constructing Table 5

Table 5 contains information regarding the adverse events reported in each study and the method the authors used to ascertain the adverse events. Studies with levels of evidence of I-V that reported adverse events are included in this table. As in Table 2, group design and single subject design studies are in separate sections of the table. In both sections, list studies in chronological order. Report the method used to ascertain adverse effects (e.g., not stated; longitudinal registry in a defined sample or population; active, systematic surveillance). If another method was used, please describe. Indicate the adverse events reported in the studies and the number of cases associated with each adverse event. All abbreviations used in the table are noted at the bottom of the table in order of appearance in the table. If no adverse events are reported in the studies reviewed, Table 5 is omitted and the lack of reported adverse events is mentioned in the discussion section (question #5). In the event that an article included in Table 5 is not included in Table 2 (e.g., a case study reporting exclusively on adverse events of an intervention), identify the study with an asterisk and include a footnote under the table that indicates the study was not included in Table 2 because it only reported on adverse events.

Table 5 Reported Adverse Events

<b>Study</b>	<b>Level of evidence</b>	<b>Total n</b>	<b>Method of Ascertaining Adverse Event</b>	<b>Description of Adverse Events Reported (number of cases)</b>
<b>Group Design Studies</b>				
Lewis 1994	IV	10	ns	

ns, not stated