

**Guideline adherence in physical therapy:
a systematic review**

A.G.J. Kolman

Student nr.: 3344091

Systematic review

Utrecht University

Clinical Health Sciences

Physiotherapy Science

Supervisor: Dr. H.J.M. Hendriks

Care and Public Health Research Institute (CAPHRI)

Department of Epidemiology

Centre of Evidence Based Physiotherapy

Maastricht University

Examinator: Dr T. Takken

Medical Physiologist

Clinical Health Sciences

Program in Physical Therapy Science

Utrecht University

Eindhoven, august 2012

ONDERGETEKENDE

bevestigt hierbij dat de onderhavige verhandeling mag worden geraadpleegd en vrij mag worden gefotokopieerd. Bij het citeren moet steeds de titel en de auteur van de verhandeling worden vermeld.

Met dank aan Diederik Bolkestein

Content	Page
Abstract	1
Background	2
Methods	3
Results	4
- Objects of study	5
- Study methods	6
- Figures provided for PT's solely or combined with other health care workers	7
- Assessment, classification and scoring methods	7
- Report methods	8
- Adherence and/or current practice, the use of evidence based guidelines, validation, operationalization and reliability	8
Discussion	9
Conclusion	13
Figure 1: Flow-chart of selection and retrieval process	15
Table 1: Methods used for assessing and classifying adherence	16
References	21

Abstract

Background

The development and use of evidence based clinical practice guidelines is widely advocated. Adherence with guidelines seems however to be moderate. To date no review has been published addressing the latest state of the art of adherence with clinical practice state of the art guidelines in physical therapy or even with non-evidence based statements, protocols or treatment methods. The aim of this systematic review is to establish the contemporary stand of adherence in physical therapy with these guidelines, to compare the adherence rates and the methods used to assess adherence and to assess whether and how validity and reliability of the assessment instruments are established.

Methods

A literature search on Pubmed, Cochrane Central register of Controlled Trials, EMBASE, CINAHL, PsycINFO, PEDro, ERIC, Health Psychosocial Instrument (HaPI), Social Sciences Citation Index (SSCI), Science Citation Index Expanded, BMJ Clinical Evidence and Google. Inclusion criteria were 1) participants had to be PT's and 2) studies had to be written in English, German or Dutch.

Results

Out of 350 references 72 studies were included. Studies varied concerning study methods and adherence assessment methods used, the way adherence was classified i.e. what was considered to be adherence, the adherence scoring or rating method, the type of performance activities or the aspects of care, the kind of activity or if concerning diagnostic or treatment activities what condition, aspects of care, guideline or treatment method was addressed. Next to that, there were differences in reporting and whether items were operationalized, validated and /or reliability was tested. Adherence rates varied, ranging from 1% to 100%.

Conclusions

The differences of methods used for assessing and reporting adherence makes a fair comparison impossible and highlights the lack of a uniform method to do so. These differences provide substance for recommendations to create such methods.

Background

For improving quality of healthcare the development and use of evidence based clinical practise guidelines is widely advocated (1). These guidelines provide practitioners opportunities to systematically apply scientific evidence in practice (2-6), to improve and monitor quality and efficiency of practitioners' performance (3,6,7), to enhance transparency of practice (7) to decrease variability in professional practice and to legitimise their profession in the eyes of external stakeholders (1). Within physical therapy the Royal Dutch Society for Physical Therapy has published 18 evidence-based guidelines up to 2007 as part of a national quality assurance program (8). Also, a series of evidence-based statements outlining the efficacy of physiotherapy for management of conditions, amongst them neck pain (9) and low back pain (10) were produced by the Australian Physiotherapy Association (8). Moreover, to promote collaboration the World Confederation for Physical Therapy has prioritised the development and implementation of guidelines to facilitate evidence-based practice (11-13).

In order for guidelines to be beneficial they need to be implemented and professionals need to adhere to the guidelines (14,15). Systematic reviews concerning the effectiveness of implementation strategies used for clinical guidelines showed only small to moderate effects for different interventions like reminders (16), dissemination of educational materials (16), and audit and feedback (16,17), educational outreach (16,18), the use of opinion leaders (14) and educational meetings and workshops (19). A recent systematic review concerning implementation strategies of physiotherapy clinical guidelines showed comparable results implying that multifaceted interventions based on educational meetings to increase implementation of clinical guidelines may improve some outcomes of professional practice but do not improve patient health or reduce cost of care(8). Also, adherence with clinical guidelines seems to be moderate (20,21).

Although much is known about the effectiveness of implementation strategies to date no review has been published addressing the latest stand of adherence with clinical practise guidelines in physical therapy or even with non-evidence based statements, protocols or treatment methods. Therefore the primary aim of this systematic review is to establish the contemporary stand of adherence with clinical practise guidelines and non-evidence based statements, protocols or treatment methods in physical therapy and compare the adherence rates.

Adherence can be assessed in different ways resulting in different levels of adherence. For a short but comprehensive overview concerning the advantages and disadvantages of each method see Rutten (2006) (22). For the purpose of a fair comparison of adherence with different clinical guidelines in physical therapy the methods of assessing adherence in each of the included studies is

taken into account. Also, as with all instruments, adherence instruments need to be valid and reliable (23). Whether and how validity and reliability are established is also addressed in this study.

Methods

A literature search was conducted on Pubmed, Cochrane Central register of Controlled Trials, EMBASE, CINAHL, PsycINFO, PEDro, ERIC, Health Psychosocial Instrument (HaPI), Social Sciences Citation Index (SSCI), Science Citation Index Expanded, BMJ Clinical Evidence and Google.

At the onset of this study there was no indication about the number of studies expected to be found but it was thought to be not more than around 30 altogether. Therefore a broad search strategy was set up, there were no explicit exclusion criteria formulated and the inclusion criteria were set with the only limitations that physiotherapists (PT's) had to be included and studies were written in English, German or Dutch. Early non-published explorations and preparatory searches revealed that several different expressions are in use to address the concept of adherence by health providers. Within the quality paradigm adherence belongs to the 'process' of care. Medical researchers and professionals prefer to use the term 'appropriateness' of care. Although there seems to be a huge overlap in the research fields the distinction in use is remarkable. When using the term process a lot of research performed by medical researchers is not found, when using the term appropriateness a lot of research concerning the quality paradigm is not found. Others use the term adherence or it's synonyms like compliance, concordance or, for instance within the field of psychology, fidelity. In nursing the terms research utilization or knowledge utilization are common. In order to yield a broad search the following MeSH-terms and key words were used: physical therapy, guideline adherence, adherence, concordance, compliance, fidelity, appropriateness of care, process of care, quality indicators, guidelines, protocol, current practice. The term current practice was used to search for non-evidence based statements, protocols or treatment methods. Next to that, the terms 'benchmark' and 'employee performance appraisal' were initially used. This search yielded over 1700 references with almost no overlap. Since this was thought to be too much these terms were omitted in a new final search.

The search was performed by one person (AK). Screening and selecting of the titles and abstracts of the articles found by this search was done independently by two researchers (AK, DB). Screening of the full text of the selected articles was again done by one person (AK). In addition, references of the selected studies were screened for additional relevant studies by one person (AK). There was no restriction in study types being included. In case of randomised trials, controlled trials, controlled before and after studies, or interrupted time series studies investigating the implementation of clinical guidelines by PT's the different groups and/or baseline measurements were regarded as a

cohort. Methodological quality was not assessed since at the time of the analyses no instrument for assessing the methodological quality of the studies was known to the researchers.

Results

The search through to June 26, 2009 yielded 1164 references. Only searches in Pubmed, Cochrane Central register of Controlled Trials, EMBASE and CINAHL yielded references using the full broad search. The search strategy and results are provided in additional file 1. After removing duplicates 350 references remained for screening. Screening of titles and abstracts by two reviewers left 75 references. After full-text screening 47 references concerning 43 studies remained. For one study two articles (24,25) and for another study four articles (26-29) were found. One review (8) concerning effectiveness of implementation strategies contained three studies (24,30,31), two of which were already found (24,30) and the third was subsequently included (31). Eleven studies could not be retrieved. Ten of these were provided by the authors. Twenty references were excluded because adherence was not measured. Other reasons for exclusion were: being a study protocol (2 studies), no PT's included (3), concerning the assessment of practice skills of participants in an exam situation (1), concerning a visitation project in which personal recommendations were provided (1). Through screening of the references of the included articles another 27 eligible studies were found. Two of these studies concerning respiratory and mobilization therapists (32) respectively manual therapists were(33) included since these therapists are mostly (specialized) PT's. Also, three studies (34-36) which concerned adherence were not found through the search nor the reference screening but were already in the possession of the first author. The latter three articles and the 27 studies found by screening were not further screened for references. Finally, 73 studies were included. In [figure 1](#) the selection process is depicted.

The first study addressing adherence was performed in 1990. During the nineties 9 more studies were performed. The other studies were performed after the change of the millennium with 18 studies being performed between 2000 and 2004 and the other 46 after 2004.

Studies varied concerning study methods and adherence assessment methods used, the way adherence was classified i.e. what was considered to be adherence, the adherence scoring or rating method, the type of performance activities or the aspects of care, the kind of activity or if concerning diagnostic or treatment activities what condition, aspects of care, guideline or treatment method was addressed. Also, adherence rates varied, ranging from 1% to 100%. Additional file 2 provides a table with comprehensive analyses.

Objects of study

There is a myriad of activities, conditions, aspects of care, guidelines or treatment methods in which adherence was assessed. Twenty-two studies (22,24,31,33,37-54) addressed low back pain. Nine of these studies (22,24,38,39,41,43,44,46,49,54) used official guidelines of which five used the Dutch guidelines for (nonspecific) low back pain (22,24,44,46,54), one measured PT's adherence to the Dutch Low-Back Pain Guideline for General Practitioners (38) and one concerned the assessment of red flags in patients with low back pain (49). In one study (37) both management of low back pain and management following total knee replacement were addressed. In 11 studies (26,32,55-63) hand washing was studied. One of these studies (56) aimed at isolation precautions such as hand washing and barrier precautions. Six studies concerned stroke (34,64-68). Four of which used official guidelines (64,65,67,68) and two studied the Bobath concept (34,66). One of these studies solely addressed the use of outcome measures recommended in the guideline (67). Four studies (69-72) surveyed osteoarthritis. Three of the knee (70-72) and one of the hip and the knee (69). Also, two of these studies (69,72) used official guidelines. Another four studies (73-76) addressed the use of safety measures whilst applying diathermy. Four studies (15,77-79) addressed ankle sprain of which two used the Dutch Clinical Guideline Acute Ankle Injury (15,79) and another one used the same guideline amongst assumptions on expected treatment such as goals and interventions that are derived from the literature (77). Two studies (80,81) concerned adherence to the American Association of Cardiopulmonary Rehabilitation established minimum and preferred guidelines for cardiac rehabilitation personnel regarding educational degree attainment, licensure, and certification. Further studies concerned The Dutch clinical practice guideline for cardiac Rehabilitation (82), the Australian Physiotherapy Association's Protocol for Pre-manipulative Testing of the Cervical Spine (83), The Australian Clinical Practice Guidelines for Whiplash (30), the use of standardized outcome measures included in the Victorian (Australia) Transport Accident Commission (TAC) Clinical Justification Model (84), chronic shoulder complaints (85), service provision of emergency on-call respiratory physiotherapy (86), provision of 24 hour access to PT's and radiographic services in chest physiotherapy (87), sudden muscle pain appeared during sporting activity (88), upper extremity musculoskeletal disorders (89), The Dutch Guideline for Physiotherapy in Parkinson's Disease (90), prevention of hip fractures after falls (91), sputum clearance and the periodical assessment of patients suffering from cystic fibrosis (92), assessing compliance rates for following three electromyography clinical practice guidelines of carpal tunnel syndrome, ulnar neuropathy at the elbow and cervical radiculopathy (93), referral to orthopedic surgeons for surgical intervention by extended scope PT's (94), goal setting and making prognosis (95), involving patients

in goal setting (96), reporting of a suspected case of elder abuse to proper legal authorities (97), the combined assessment of non-specific neck pain and the proposed screening process for direct access to physical therapy in the Netherlands (36), quality of physiotherapy patient documentation using the Dutch Clinical Practice Guideline on Physiotherapy Documentation (35) and assessing student performance of entry level of professional skills by clinical instructors (98).

Study methods

A questionnaire survey was used by 19 studies (30,38,44,45,53,54,67,72,74,76,80,81,83,86-88,91,92,97) one of which questioned patients concerning PT's' adherence (92) and one of which used a telephone interview (88). In all these studies respondents were questioned about their performance and adherence was subsequently determined by the researchers. In nine studies (44,54,66,67,79,84,85,90,98) self-reported adherence was assessed in which respondents themselves rated the extent to which they adhered to the prescriptions, protocol, recommendations or guidelines. In three studies (44,54,67) both self-reported use as well as self-reported adherence were measured. Nine vignette studies were performed accompanied by a questionnaire (22,34,39,41,50,52,54,70,78) and one vignette study (45) was performed accompanied by a telephone interview. Also, two studies (57,59) used clinical scenarios accompanied by a questionnaire. One study (31) used a questionnaire survey before and after a RCT and one study (85) used a post-intervention questionnaire in an RCT in which both PT's (self-report) and patients were questioned concerning PT's' adherence. Also one study (34) took place as an intervention check for a large outcome study measuring the effects of Bobath-based therapy. A quasi-experimental study combined with observations was performed by five studies (26,32,55,61,62). An observational survey was used by six studies (33,56,58,63,75,76) of which one study (75) concerned an assessment of the environment and one study (33) in combination with an interview. Twelve studies (35,37,42,48,49,51,64,65,68,93-95) performed a retrospective chart review of which one study (94) used letters sent by the orthopedic surgeons of the patients involved after consultation. Twelve studies (15,30,36,40,43,47,69,71,73,77,82,83) performed a prospective registration survey and one study (46) a retrospective registration survey in which existing registration forms were analyzed. Next to that, a prospective registration survey was used in the intervention arm of an RCT in four studies (24,30,85,89) of which in two studies (24,30) also the control arm was surveyed, in an interventional randomized cohort study (60) with four study phases and in a non-randomized non-controlled intervention (quasi-experimental) study (84). One study

(22) used data from a previous study (24) to compare with new data. Also, one study (96) used audiotapes.

Figures provided for PT's solely or combined with other health care workers

In ten studies figures were not provided for PT's solely but together with other health care workers (HCW's) as GP's (50), care assistants (26), technicians (63), occupational therapists (OT's) or team (64), OT's, nurses, nursing assistants and doctors (91), 'others from several allied health fields' (80), nurses and radiologic technicians (58), EKG (electrocardiograph) technicians, anesthesia technicians and housekeeper (56), midwives, respiratory and mobilization therapists, radiology technicians, nutrition therapists, as well as HCW's of all professional categories apart from nurses, nursing assistants and doctors (32), OT's, phlebotomists, dieticians etc. (61). All the other studies provided numbers per profession or at least solely for PT's.

Assessment, classification and scoring methods

In 52 studies adherence was assessed and classified dichotomously (yes/no). Fulfilling an item or indicator was considered to be adherence and not fulfilling an item or indicator was considered to be non-adherence. Three studies used a questionnaire with an ordinal (66,90) or interval (84) scale to assess adherence but dichotomized item scores in the analyses to classify adherence although one of these studies (84) did not use the dichotomization furthermore in the analyses. Two studies (61,83) used both a dichotomized method and ordinal scales. In a further 7 studies (35,50,57,59,71,79,85) adherence was assessed and classified using an ordinal level of measurement, 4 more studies (22,36,89,98) used an interval level of measurement and 3 studies (45,54,93) used both ordinal and interval scales.

One study (53) used ranking, in one study (52) the methods were not provided and in one study (38) the methods could not be provided in terms like dichotomously, ordinal or interval because the applied interventions and the recommendations were compared descriptively. In the studies in which adherence was assessed and classified using ordinal or interval scales a myriad of combined methods to assess and classify adherence were used. Mostly semantic combinations or 5- or 6-point scales were combined with weighted scores added up, non-weighted scores added up, proportions, ranges of proportions, risk levels or levels of evidence. [Table 1](#), which is a summarized version of the comprehensive table in additional file 2, provides a descriptive overview of the studies in which adherence was assessed and classified using ordinal and interval scales.

In most studies the figures were provided per single recommendation (additional file 2) in which each recommended intervention was applied by a group of PT's to a group of patients. In some

studies also an overall rate was calculated for all recommendations together, either by adding up the percentages and dividing this by the number of items or by appointing scores to each item and adding these up. In one study (93) the overall calculation per complaint was based on weighted items which were awarded weights of 2, 4 or 6 maximum points. In another study (98) each of the ten items were appointed a score from 1 to 5. Subsequently, each PT got a score between 1 and 50 and the figures were provided by the percentages of PT's who scored between 41-50, 31-40 and <30. One study (22) used the same seven questions for three vignettes in which each question was assigned a weighted numeric score. Per vignette, this score depended on the specific recommendations the guideline provided for the particular case described. In the same study also recording forms were used and each of them got an overall score for all four recommendations. In yet another study (68) each of the 38 recommendations was appointed a score of 1 point when the recommendation was met. If all 38 recommendations were met, there was a perfect score of 38. Also, one study (46) also used an overall score of therapist who treated at least five patients and five studies (22,24,46,84,93) used an algorithm.

Report methods

In 11 studies (38-42,51,53,70,72,83,87) there was reflection on adherence to recommendations in the discussion. Seven of these studies (38,39,41,42,53,70,83) explicitly stated that the findings were compared with existing recommendations. In the other studies adherence was reported in the results section.

Adherence and/or current practice, the use of evidence based guidelines, validation, operationalization and reliability

Forty-three studies solely addressed adherence. The other 30 studies addressed both current practice and adherence. Studies differ considerably whether or not validated and operationalized measures were used and the way this is reported. Five studies (35,36,82,86,90) stated they used a systematic approach or formalized method to choose, develop and validate indicators to assess adherence. Four of these studies (35,36,82,90) used criteria for the indicators as impact on outcome and societal cost (82), effectiveness, efficiency, acceptability and measurability (90), being representative for the 6-phases of the structured physical therapy assessment, room for improvement and being measurable and broadly applicable (35) and acceptability, feasibility and sensitivity to change(36). Neeleman-Van der Steen (2008) (82) and Nijkrake (2009) (90) rated these criteria using a 4-point respectively a 5-point Likert scale. Twelve of the studies (31,33,38-42,51,70,72,74,78) addressing both current practice and adherence stated they have only used

validated measures to assess current practice, two studies (44,84) only to assess adherence, 11 studies (37,45,50,53,66,71,76,81,83,86,97) to assess both current practice and adherence and five studies (43,47,77,87,92) didn't state they have used any validated measures at all. In total 32 studies stated they used validated measures to assess adherence. Fifty-three studies, of which 27 studies didn't state they used validated measures to assess adherence, stated they used evidence based guidelines or provided the importance of the process being measured. For instance, Saliga (2004) (97) studied reporting suspected elder abuse according to the law and provided information about the importance of this law. Also, many of the studies concerning hand washing and the studies concerning safety measures during diathermy emphasized their importance. One study (83) explicitly stated that there was a lack of research underpinning the protocol or method. Also, in one study (46) the notion of the authors to use explicit operationalized and valid criteria was stated but it was not provided whether this was done as such. Also, many studies explicitly stated they have used operationalized measures whereas others didn't state this but provided clear measures in either tables, text or appendixes. Again others didn't state nor provide anything about operationalization. Seven studies (35,36,49,56,93,98) stated they established reliability of which one study (36) didn't provide figures. Additional file 3 provides an overview.

Discussion

The primary aim of this systematic review to establish the contemporary state of the art of adherence with clinical practise guidelines and non-evidence based protocols in physical therapy is met that far that only a general overview could be provided. Due to a plethora of assessment methods, classification methods, scoring methods, numbers of items, aspects of care, the type and number of HCW's involved and the way numbers of HCW's are provided any comparison of the adherence rates turned out to be impossible.

The 43 selected studies were initially found to be very much and it was tried to reduce this number by omitting for instance all studies which did not address diagnosing or treating or which did not use solely PT's. However, this would leave out interesting studies concerning diagnosing and treating and concerning used methods. It was therefore decided to include the 43 studies and use these for screening the reference lists.

Due to leaving out the terms 'benchmark' and "employee performance appraisal' an unknown number of eligible studies were probably not found. Also, other eligible studies in which adherence was assessed but not explicitly mentioned as such were not found. For instance, Kwakkel (1999) (99) reported that the patients were given the recommended five treatment sessions a week. It's very

well possible that much more studies do address adherence in such an implicit way. These studies will have been missed by our search .

The differences found in this study in the ways adherence was assessed, classified and scored, whether or not validated and/or operationalized measures were used and the ways the results are reported highlights the lack of a uniform method in assessing and reporting adherence. Due to this deficiency researchers were more or less forced to create their own methods. Especially concerning assessing, classifying and scoring adherence as demonstrated in [table 1](#). The lack in using uniform methods could be reasonably explained by the fact that it's a fairly new topic in research in healthcare and that it has, in our opinion, been a kind of by product of or at least of less interest with respect to the search for evidence, the development of guidelines and the implementation of these guidelines. An underlying reason for this could be that measuring adherence demands a fair degree of introspectiveness for the profession as a whole as well as for individual professionals. This can be anticipated by the respondents as a threat and by the assessors as a kind of judging or an interference with someone else's work. Berwick (1989) (100) already stated that: 'In fact, practically no system of measurement – at least none that measures peoples performance – is robust enough to survive the fear of those who are measured'. The lack in uniform reporting of adherence is shown by the 11 studies (38-42,51,53,70,72,83,87) which assessed both current practice and adherence in which current practice was addressed in the results section and adherence in the discussion. All these studies reflected on adherence in the discussion of which seven studies (38,39,41,42,53,70,83) explicitly stated that the findings were compared with existing recommendations. Although, it's common in research to reflect in the discussion on the results it could be argued to recommend to report adherence to prescriptions in the results section. Only very recently there is attention for better methods to asses adherence . The five studies (35,36,82,86,90) which state they used a systematic approach or formalized method to choose, develop and validate indicators to assess adherence are relatively new. The earliest stems from 2006 and the literature they refer to (101,102) goes back to 2001. In the other studies researchers used their own methods of which the methods of validating the instruments, when validated at all, were commonly used methods as using pilots and experts.

The attention for better methods to assess adherence can even be regarded as a next logical step in the development process of the quality of healthcare from the search for evidence through the development of guidelines and the implementation of these guidelines. After the evidence was established the next step was to provide this evidence to healthcare workers in a concise way, immediately raising the problem of dissemination and implementation and subsequently the issue of

how to establish the effects of the dissemination and implementation. The latter leading directly to the topic of methods to assess adherence.

As for all measures, for measures assessing adherence validity and reliability need to be determined (102-105). In our results the number of studies which stated they established reliability was a mere seven studies (35,36,49,56,60,93,98). We could find no explanation why this number is so low. Concerning the validation of the measures the notable differences between studies assessing both current practice and adherence raises the attention to the differences between the concept of current practice and the concept of adherence. Current practice refers to what is commonly performed in daily practice. An instrument to assess current practice for the treatment of a certain complaint should therefore reflect all possible diagnostic activities and interventions or modalities. According to Webster's Students Dictionary (106) to adhere to means 'to follow closely or without deviation'. According to Merriam-Webster's Learner's Dictionary (107) it means 'to act in the way that is required by (something, such as a rule, belief, or promise)'. According to the Oxford English Dictionary (108) it means 'to cleave to an opinion, practice, or method'. In the medical and allied healthcare professions to adhere to means 'to follow closely or without deviation to what' or 'to cleave to what' or 'to act in the way that' is anticipated i.e. believed as good practice, whether it's evidence based or not. This anticipated good practice is often provided in prescriptions as protocols, standards, algorithms, recommendations, guidelines or methods. Therefore, an instrument to assess adherence should, in contrast to an instrument to assess current practice, reflect only diagnostic activities and interventions or modalities which stem from prescriptions. In case both current practice and adherence are assessed validated measures for both concepts need to be used or if one measure is used it should be validated for both current practice and adherence.

Although 27 studies didn't state they used validated measures to assess adherence they did state they used evidence based guidelines or provided the importance of the process being measured. This states a certain implicit content validity of the used measures with these guidelines or processes. The importance was demonstrated by explaining why the performance is undoubtedly anticipated as good care. It's obvious and/or commonly accepted that reporting suspected elder abuse, hand washing and taking safety measures are performances of good care. This implicit content validity can partly be explained by the above mentioned meanings of the words adherence, prescription, standard, protocol, recommendation and guideline. Also, some authors, for instance Bekkering (2005) (24) did not validate the criteria used for assessing adherence but did operationalize these criteria to clarify the conditions under which the criteria are considered to reflect adherence when fulfilled. This operationalization can be considered as a kind of validation since it has been well deliberated whether the criteria did measure what was supposed to be measured. Next to that,

some authors, like Van Peppen (2008) (67), used guidelines which were already clearly operationalized. Van Peppen (2008) (67) used the guideline which states that recommended core outcome measures should be used at specific time points which are given in the guideline. This guideline needed no further operationalization nor validation.

The distinction between studies explicitly stating that measures were validated and other studies providing content validity in an implicit manner again raises the attention to the lack of a uniform method for reporting. It's suggested here to either agree upon always explicitly stating that the measures were validated or to agree upon that a) mentioning that evidence based measures were used or b) providing a thorough explanation concerning the importance of the process being measured or c) providing the operationalization process and and/or clear operationalized indicators are also sufficient in reporting validity.

Next to the implicit content validity, measures build on evidence based guidelines or evidence taken directly from literature do have construct validity since the relationship between the process, e.g. the intervention, with the outcomes is established (105,109). However, since the evidence is not always very strong (109) and experts often disagree on the interpretation of evidence(102) it's recommended to validate and operationalize measures by combining evidence with consensus methods (102,109). On the other hand, many guidelines and recommendations are developed using thorough literature searches and consensus methods. So why bother to repeat a consensus method. Rather, it will be more efficient to immediately operationalize recommendations or guidelines in measurable indicators whilst they are being developed.

It's also recommended to use and rate criteria which the measures or indicators have to meet (102,110) which is done by four of the five studies (35,36,82,86,90) which stated they used a systematic approach or formalized method to choose, develop and validate indicators to assess adherence. These four studies (35,36,82,90) used criteria for the indicators as impact on outcome and societal cost (82), effectiveness, efficiency, acceptability and measurability (90), being representative for the 6-phases of the structured physical therapy assessment, room for improvement, being measurable and broadly applicable (35) and acceptability, feasibility and sensitivity to change (36). The use of these criteria in these studies stem from studies outside the field of physiotherapy which have proposed similar and other criteria like relevance to clinical performance (105,111), societal cost (105), level of evidence (103,105,112), feasibility (102,112,113), impact on outcome (105,113,114) measurability (105,114), room for improvement (103,105,113), plausibility (113), acceptability (102), sensitivity to change (102) and overall utility (113). Although measurability seemed difficult to appraise by experts in a study to implement the Dutch infertility guidelines (115). There seems, however, not to be any consensus on which criterion

should be used when or be used at all. It would therefore be advisable to agree upon one set of criteria to which measures have to meet.

For both the rating of the validity as well as the separate criteria specific scales can be used as done by two of the included studies which used a 4-point (82) and a 5-point (90) Likert scale. Studies outside the field of physiotherapy also used a 5-point scale (105,113) or a 9-point scale (111,113). The latter stems from the RAND/University of California, Los Angeles–modified Delphi method (112). The five studies (35,36,82,86,90) which stated they used a systematic approach or formalized method to choose, develop and validate indicators to assess adherence used quite similar methods although they differed in detail concerning the number and sequence of steps taken and the use of criteria to which the indicators have to be met. They preceded the recently developed AIRE Instrument for Appraisal of Indicators through Research and Evaluation (116). An instrument in which the quality of all indicators of a set of indicators is appraised separately by use of 20 criteria grouped into four quality domains: 1. purpose, relevance and organizational context, 2. stakeholder involvement, 3. scientific evidence, 4. additional evidence, formulation and usage. This instrument is more comprehensive than the systematic approaches or formalized methods mentioned above and includes some of the criteria mentioned above i.e. impact on outcome, level of evidence, room for improvement. It does not state however why these were chosen and whether or not other criteria were screened and omitted.

Another notable distinction is that in self-reported use and self-reported adherence. In self-reported use the respondent is not asked whether she or he adheres to the prescriptions, protocols, recommendations or guidelines. The decision whether or not she or adheres is made by the assessor. When, however, the respondent is asked directly whether she or he adheres to the prescriptions, protocols, recommendations or guidelines this decision is made explicitly by the respondent her- or himself. Some studies used questions in which the word 'use' is combined with phrases like 'according to the recommendations'. In this case the respondent can immediately know that adherence is assessed and therefore she or he decides implicitly her- or himself whether she or he adheres. It's recommended here that when using self-report it should be made clear whether it's self-reported use or self-reported adherence and who decides what's considered to be adherence.

Conclusion

The primary aim of this systematic review to establish the contemporary state of the art of adherence with clinical practice guidelines and non-evidence based protocols in physical therapy is met that far that only a general overview is provided. Due to a plethora of assessment methods, classification methods, scoring methods, numbers of items, aspects of care, the type and number of

HCW's involved and the way numbers of HCW's are provided any comparison of the adherence rates turned out to be impossible. The differences found in this study in the ways adherence was assessed, classified and scored, whether or not validated and/or operationalized measures were used and the ways the results are reported highlights the lack of a uniform method in assessing and reporting adherence. The same differences provide substance for recommendations to create such methods.

Figure 1: Flow-chart of selection and retrieval process

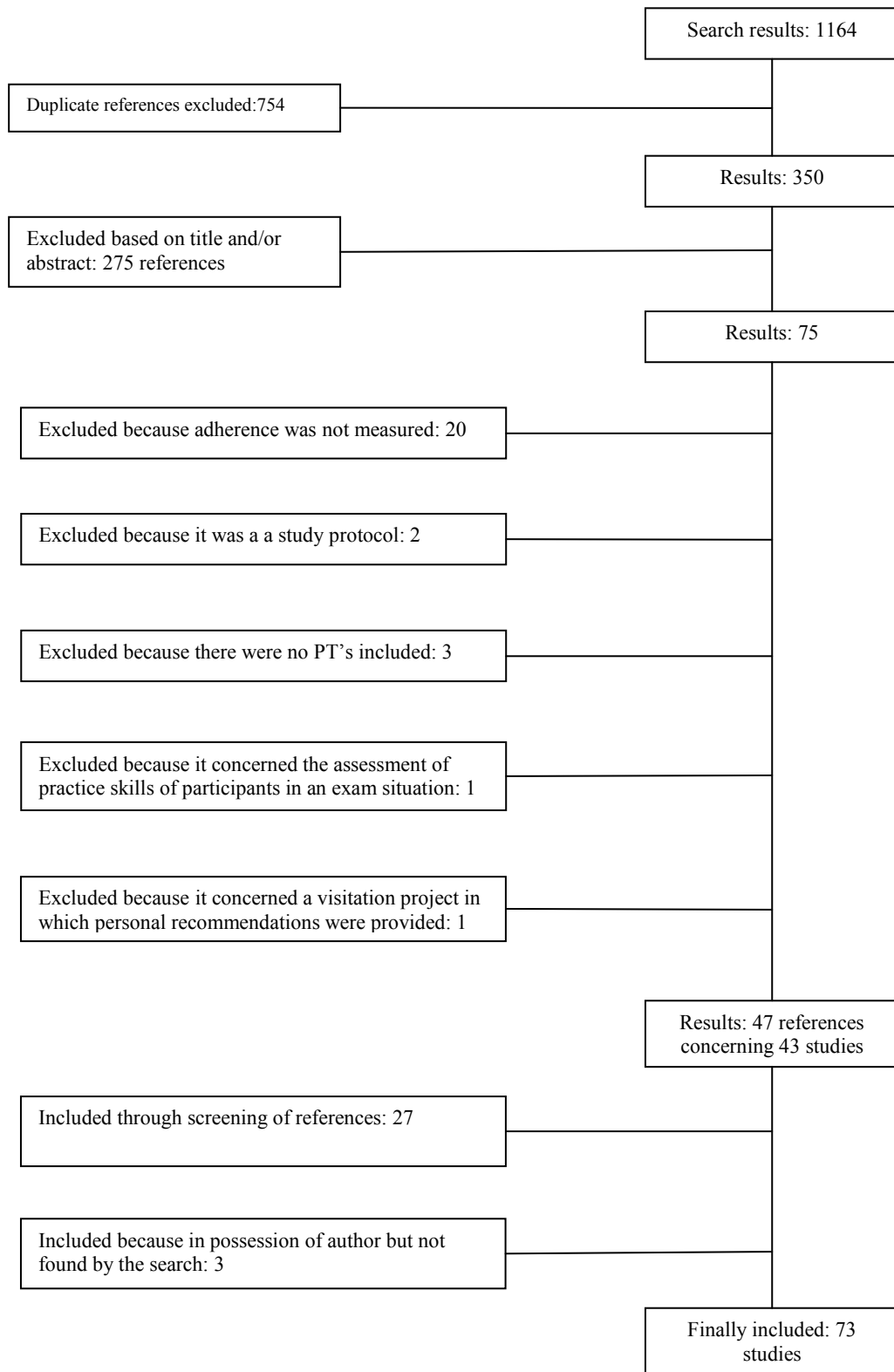


Table 1: Methods used for assessing and classifying adherence

Author, year (country)	Adherence assessment method.	How was adherence classified?
Alexander, 1997 (US).	Nineteen specific clinical scenarios using a 6-point scale. The scale ranged from 0 to 100% (0, 20, 40, 60, 80, 100%), corresponding to how often the respondent washed his or her hands in the specific situation.	Using an ordinal scale. A score of 6 indicated an acceptable response, 1 indicated an unacceptable response according to the authors interpretation of CDC guidelines. The scores in between 2-5 signified a range between acceptable and unacceptable, indicating an inconsistency in practice. In the results section the acceptable compliance score was set at 5 and above.
Hrachovy, 2000 (US).	Questionnaire using a 5-point Likert scale: 1. never 2. rarely 3. occasionally 4. frequently 5. always.	Using an interval scale. The ratings on the 5-point Likert scale reflect directly to the level of adherence in which 'never' reflects non-adherence and 'always' reflects full adherence. Moreover, the qualifications 'always', 'frequently', 'occasionally', 'rarely' and 'never' were quantified ranging from 1. never till 5. always. For the 10 items of the questionnaire the points were added up ranging from 1 to 50. These scores were categorized in three categories in which scores 41-50 means very consistent adherence, scores 31-40 means moderately consistent adherence and scores < 30 means inconsistent adherence.
Devine, 2002 (US).	Nineteen specific clinical scenarios designed into questions (statements) using a six-point scale. The scale represented an estimate of the percentage of hand washing occurrences (0, 20, 40, 60, 80 100%) necessary following each clinical scenario.	Using an ordinal scale. A score of 1 was given for noncompliance and a score of 6 was given for compliance according to the authors interpretation of CDC guidelines for each clinical scenario presented. Scores from 2 - 5 indicated a range of compliance and therefore inconsistency in hand washing behaviors.
Magaray, 2004 (AU).	By asking a direct question whether information about the dangers is given using answering options: 1. always 2. sometimes and whether consent is asked using answering options: 1. only at the first occasion 2. at every occasion 3. for every High Velocity Thrust technique 3. not with every patient 4. never.	Using an ordinal scale. Concerning provision of information the answers 'always' and 'sometimes' reflect a classification in degrees of adherence in which 'always' was considered to be adherence. Concerning seeking consent the degree of adherence is operationalized depending on the patient and certain occasions as first or every occasion or the occasion in which the HVT technique is used. It's not clear what was exactly considered to be adherence but 'never' is considered to be non-adherence.
Leemrijse, 2005 (NL).	1. By ways of a direct question asking PT's to which extent they treated their patients according to the guidelines, using a 6-point-	Using an ordinal scale. For the direct question the ratings on the 5-point scale reflect directly to the level of adherence in

	<p>scale:</p> <ol style="list-style-type: none"> 1. never 2. < 10% of the patients 3. 10–50% of the patients 4. 50–90% of the patients 5. > 90% of the patients 6. always. <p>2. Questions asked whether PT's treated their patients according to individual recommendations in the guidelines:</p> <ol style="list-style-type: none"> 1. yes 2. sometimes 3. no. 	<p>which 'never' reflects non-adherence and 'always' reflects full adherence. Also, concerning the individual recommendations the answers 'yes', 'sometimes' and 'no' reflect a certain degree of adherence in which 'yes' reflects full adherence and 'no' reflects non-adherence.</p>
Mikhail, 2005 (CA).	<p>By using:</p> <ol style="list-style-type: none"> 1. A clinical vignette in which each intervention was reported according to its evidence of effectiveness (strong, moderate, limited, or none) 2. A checklist of 13 interventions with known evidence of effectiveness by rating the frequency of use on a 5-point scale ranging from 'never' to 'always' in relation to a client similar to the one in the vignette. 	<ol style="list-style-type: none"> 1. Using an ordinal scale. Interventions were classified according to the levels of evidence of effectiveness in which the highest level of evidence was considered to be highest degree of adherence. 2. Using an interval scale. The ratings on the 5-point Likert scale reflect directly to the level of adherence in which 'never' reflects non-adherence and 'always' reflects full adherence.
Abrams, 2006 (AU).	<p>Using a questionnaire asking the physiotherapists "In what proportion of cases when the test could be used do you use the test?" with possible responses:</p> <ol style="list-style-type: none"> 1. don't know the test 2. never 0% 3. sometimes 1–30% 4. frequently 31–60% 5. always 61–100%. 	<p>Using an interval scale. In the figures only proportions are given in which a high percentage reflects a high level of adherence. It was stated that in the analyses the response categories were dichotomized. The categories 'sometimes', 'frequently' and 'always' were collapsed for comparison of change in reported use and the categories 'frequently' and 'always' to examine change in common practice. These categories were not used furthermore as such.</p>
Geraets, 2006 (NL).	<p>Questionnaires using a 5-point ordinal scale</p> <ol style="list-style-type: none"> 1. fully convinced. 2. strongly convinced. 3. reasonably convinced 4. little convinced. 5. not at all convinced. 	<p>Using an ordinal scale concerning GET by using the terms fully, strongly, little and not at all in combination with convinced. Not at all convinced was considered to be non-adherence. Fully convinced was considered to be full adherence.</p>
Meijer, 2006 (NL).	<p>By marking all activities performed and/or discussed on a checklist after every session. These activities were later classified into physical and psychological components, return-to-work, relaxation and energy components. The percentage of the planned sessions of the major components that actually took place was calculated.</p>	<p>Using an interval scale. Compliance was determined to be good if more than 74% of the scheduled sessions took place, moderate between 50%–74% and poor if less than 50% of the treatment protocol was carried out.</p>

Oostendorp, 2006 (NL).	A quality evaluation instrument consisting of 44 indicators being scored from 0 to 3: 0=absent 1=somewhat present 2=mostly present 3=completely present na=not appropriate.	Using an ordinal scale. The scores reflect directly to the level of adherence in which 'absent' reflects non-adherence and 'completely present' reflects full adherence.
Randle, 2006 (GB).	Using a tool for hand hygiene audits. Frequencies of scores were used with '0' being scored for a hand hygiene opportunity (i.e. '0' for opportunity) and '1' being scored for an observation of hand hygiene actually taking place.	Using an ordinal scale. Opportunities for hand hygiene were stratified into three categories: high, medium and low risk in which high risk was considered more important to be adhered to. However, these categories were not used furthermore as such but incorporated in the figures.
Rutten, 2006 (NL).	By using: 1. three clinical vignettes with 7 criteria to assess adherence to 6 recommendations . Per vignette each criterion was given a weighted score depending on the specific recommendations the guideline provided for the particular case described 2. recording forms and an algorithm based on 4 main recommendations which were operationalized in measurable performance indicators.	1. Using an interval scale. All three vignettes were supplied with the same 7 questions. To enable statistical analysis of the data, each question was assigned a weighted numeric score. Per vignette, this score depended on the specific recommendations the guideline provided for the particular case described. Answers that matched the recommendations were given the weighted scores, whereas answers that contravened the recommendations were given no points at all implicating that more relevance is given to the recommendations. A higher score was considered to be adherence. 2. Using an interval scale. Adherence to each of the four recommendations resulted in 25 points, non-adherence in 0 points. This resulted in a 0, 25, 50, 75 or 100% adherence score per recording form.
Bishop, 2007 (GB).	A vignette accompanied by clinical behaviour questions with multiple choice response options (statements) concerning advice given to patients with LBP about work (5 options), activity (4 options) and bed rest (5 options).	Using an ordinal scale. Response options were classified as being strictly, broadly or not in line with guideline recommendations.
Rutten, 2007 (NL).	A modified classification based on the scoring system used in Dutch education was used: 0-15% negligible, 16-25% little, 26-35% weak, 36-45% amply unsatisfactory, 46-55% unsatisfactory, 56-65% satisfactory, 66-75% amply satisfactory, 76-85% good, 86-95% very good, 96-100% excellent	Using an interval scale. The scores reflect directly to the level of adherence in which the dividing line between non-adherence and adherence lies between unsatisfactory and satisfactory.
Tyson, 2007 (GB).	Questionnaire using a 5-point Likert scale: 5. completely (I am a Bobath purist) 4. strongly (my work is strongly based on the Bobath concept	The questionnaire was on an ordinal level. In the analyses the scores were dichotomized. Scores of 1 – 3 were categorized as 'eclectic' and scores of 4 – 5

	<p>although I occasionally use other methods)</p> <p>3. mostly (my work is based on the Bobath concept but I do use other methods)</p> <p>2. not much (my work is based on other methods but is influenced by the Bobath concept a little)</p> <p>1. not at all (I do not use the Bobath concept)</p>	<p>were categorized as 'strongly Bobath' or Bobath 'purist'.</p>
Armantrout, 2008 (US).	<p>1. checklist with weighted items based on the strength of evidence and with diagnostic algorithms</p> <p>2. reviewer reported rating on a 7-point Likert scale</p> <p>1. absolutely incorrect</p> <p>2. moderately incorrect</p> <p>3. somewhat incorrect</p> <p>4. uncertain/neutral</p> <p>5. somewhat accurate</p> <p>6. moderately accurate</p> <p>7. absolutely accurate.</p>	<p>1. Using an interval scale. Items based on stronger evidence are given more weight assuming these items are more important to adhere to. The items are awarded weights of 2, 4 or 6 maximum points. Each item was graded from zero to the maximum points. Zero points was considered non-adherence and the maximum points were considered to be full adherence. The score for the checklist was calculated as the percentage of points acquired out of the maximum possible points.</p> <p>2. Using an ordinal scale. The terms 'incorrect' and 'accurate' in combination with the terms 'absolutely', 'moderately' and 'somewhat' reflect a classification in degrees of adherence. 'Absolutely incorrect' was awarded 1 point and was considered to be non-adherence. 'Absolutely accurate' was awarded 7 points and was considered to be full adherence.</p>
Jamtvedt, 2008 (NO).	<p>By registering actual clinical practice and analyzing the treatment modalities in three categories:</p> <p>1. "not used"</p> <p>2. "used in up to 80% of the sessions"</p> <p>3. "used in more than 80% of the sessions", which are compared to the quality of evidence on four levels: high, moderate, low or no evidence.</p>	<p>Using an ordinal scale. A combination of a classification of treatment modalities in three categories with levels of evidence was used. A high use of a high level modality was considered to be adherence and a high use of a low level modality was considered to be non-adherence.</p>
Liddle, 2009 (IE).	<p>Questionnaire in which respondents were asked to rank the type(s) of treatment they used most frequently with CLBP patients; higher ranks indicating more frequent use.</p>	<p>By ranking. A high ranking of advice and exercise was considered to be adherence.</p>
Nijkrake, 2009 (NL).	<p>Using evidence based quality process-indicators. PT's were asked to score how frequently they followed that specific process on a 5-point scale:</p> <p>0. never</p> <p>1. seldom</p> <p>2. sometimes</p> <p>3. generally</p>	<p>The questionnaire was on an ordinal level. To evaluate response distribution, item scores were dichotomized into 'frequently following the indicator' (item score 3 or 4) or 'frequently not following the indicator' (item score 0, 1 or 2). Proportions for 'frequently following indicator' were calculated for each item and the criterion for acceptable</p>

	4. always.	guideline adherence was set at > 75% of the physiotherapists frequently following an item.
Rutten, 2009 (NL).	<p>1. three clinical vignettes with 7 criteria to assess adherence to 6 recommendations . Per vignette each criterion was given a weighted score depending on the specific recommendations the guideline provided for the particular case described (see above, Rutten 2006).</p> <p>2. By one single item asking the physiotherapists to what extent they thought they were applying the guidelines for LBP:</p> <ol style="list-style-type: none"> 1. never 2. sometimes 3. regularly 4. mostly 5. always. 	<p>1. Using an interval scale. Rates from 0% to 33.3% were classified as low, those from 33.4% to 66.6% as moderate, and those from 66.7% to 100% as high adherence.</p> <p>2. Using an ordinal scale. The answering options 'never' and 'sometimes' were classified as low adherence, the option 'regularly' as moderate adherence, and the options 'mostly' and 'always' as high adherence.</p>

References

- (1) Marilyn J. Field, Kathleen N. Lohr. Guidelines for Clinical Practice: from development to use. Washington DC: National Academic Press; 1992.
- (2) H. J. M. Hendriks, H. Van Ettekoven, P. J. Van der Wees, E. R. Reitsma. Eindverslag van het project centrale richtlijnen in de fysiotherapie. Deel 1: Achtergronden en evaluatie van het project [Final Report of the National Guidelines on Physiotherapy Project. Part 1: Backgrounds and Evaluation of the Project]. Amersfoort; Utrecht: Koninklijk Nederlands Genootschap Voor Fysiotherapie, Nederlands Paramedisch Instituut, Kwaliteitsinstituut Voor Gezondheidszorg CBO; 1998.
- (3) Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999 Feb 20;318(7182):527-530.
- (4) Visser D, Elvers H, Oostendorp RAB. Vakbekwaam en wetenschappelijk verantwoord handelen, haalbaar of . . . [Acting professionally and evidence based, is it achievable or . . .]. Issue 2001;2:2-4.
- (5) Grol R. Invoeren van veranderingen in de praktijk [Introducing changes into practice]. In: Grol R, Wensing M, editors. Implementatie, effectieve verandering in de patiëntenzorg [Implementation, Effective Changes in Patient Care]. 2nd ed. Maarssen: Elsevier gezondheidszorg; 2001. p. 25-37.
- (6) Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002 Dec;11(4):358-364.
- (7) Bekkering GE, Vogels EMHM, Hendriks HJM. Evaluatie van KNGF-richtlijnen: kwaliteit, doeltreffendheid en knelpunten voor implementatie [Evaluation of KNGF guidelines: quality, efficiency and barriers for implementation]. In: Aufdemkampe G, Den Dekker JB, Van Ham I, editors. Jaarboek Fysiotherapie en Kinesitherapie Houten: Bohn, Stafleu & Van Loghum; 2000. p. 243-265.
- (8) Van der Wees PJ, Jamtvedt G, Rebbeck T, De Bie RA, Dekker J, Hendriks EJ. Multifaceted strategies may increase implementation of physiotherapy clinical guidelines: a systematic review. *Aust J Physiother* 2008;54(4):233-241.
- (9) J. Costello, G. Jull. Neck pain position statement. Melbourne: Australian Physiotherapy Association; 2000.
- (10) T. Rebbeck. Low back pain position statement. Melbourne: Australian Physiotherapy Association; 2002.
- (11) P. J. Van der Wees, J. Mead. Framework for clinical guideline development. Brussels: European Region of the World Confederation for Physical Therapy; 2004.
- (12) P. J. Van der Wees, J. Mead. Clinical guidelines 2: Developing the guidelines. WCPT Keynotes. EBP. London: World Confederation for Physical Therapy; 2006.

- (13) J. Mead, P. J. Van der Wees. Clinical Guidelines: An introduction. WCPT Keynotes. EBP. London: World Confederation for Physical Therapy; 2006.
- (14) Doumit G, Gattellari M, Grimshaw J, O'Brien MA. Local opinion leaders: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2007 Jan 24(1).
- (15) Van der Wees PJ, Hendriks EJ, Jansen MJ, Van Beers H, De Bie RA, Dekker J. Adherence to physiotherapy clinical guideline acute ankle injury and determinants of adherence: a cohort study. *BMC Musculoskelet Disord* 2007;8:45.
- (16) Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004 Feb;8(6):iii-iv, 1-72.
- (17) Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2006 Apr 19(2).
- (18) O'Brien M, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, et al. Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2007 12(4).
- (19) Forsetlund L, Bjørndal A, Rashidian A, Jamtvedt G, O'Brien M, Wolf F, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2009 06(2).
- (20) Oxman AD, Thomson MA, Davis DA, Haynes B. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *Can Med Assoc J* 1995;153(10):1423-1431.
- (21) Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud P-C, et al. Why don't physicians follow clinical practice guidelines?: A framework for improvement. *J Am Med Assoc* 1999;282(15):1458-1465.
- (22) Rutten GMJ, Harting J, Rutten STJ, Bekkering GE, Kremers SPJ. Measuring physiotherapists' guideline adherence by means of clinical vignettes: a validation study. *J Eval Clin Pract* 2006 10;12(5):491-500.
- (23) Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000 Apr 5;283(13):1715-1722.
- (24) Bekkering GE, Hendriks HJ, Van Tulder MW, Knol DL, Hoeijenbos M, Oostendorp RA, et al. Effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain: a cluster randomised controlled trial. *Qual Saf Health Care* 2005 04;14(2):107-112.
- (25) Bekkering GE, Hendriks HJM, Van Tulder MW, Koopmanschap MA, Knol DL, Oostendorp RAB, et al. Effectiviteit van een actieve implementatiestrategie van de KNGF-richtlijn 'Lage-rugpijn' [Effectiveness of an active intervention strategy for the implementation

- of the Dutch physiotherapy guideline on low back pain]. *Ned Tijdschr Fysiother* 2005 06;115(3):62-67.
- (26) Creedon SA. Healthcare workers' hand decontamination practices: compliance with recommended guidelines. *J Adv Nurs* 2005 08;51(3):208-216.
- (27) Creedon SA. Health care workers' hand decontamination practices: an Irish study. *Clin Nurs Res* 2006 02;15(1):6-26.
- (28) Creedon SA. Infection control: Behavioural issues for healthcare workers. *Clin Gov* 2006;11(4):316-325.
- (29) Creedon SA. Healthcare workers compliance with hand hygiene guidelines: findings from a quasi-experimental study in Ireland. *International Journal of Nursing Terminologies & Classifications* 2006 01;17(1):65-66.
- (30) Rebeck T, Maher CG, Refshauge KM. Evaluating two implementation strategies for whiplash guidelines in physiotherapy: a cluster randomised trial. *Aust J Physiother* 2006;52(3):165-174.
- (31) Stevenson K, Lewis M, Hay E. Does physiotherapy management of low back pain change as a result of an evidence-based educational programme? *J Eval Clin Pract* 2006 Jun;12(3):365-375.
- (32) Pittet D, Hugonnet S, Harbarth S, Mourouga P, Sauvan V, Touveneau S, et al. Effectiveness of a hospital-wide programme to improve compliance with hand hygiene. *Lancet* 2000 10/14;356(9238):1307-1312.
- (33) Strand LI, Kvale A, Råheim M, Ljunggren AE. Do Norwegian manual therapists provide management for patients with acute low back pain in accordance with clinical guidelines? *Manual Ther* 2005 02;10(1):38-43.
- (34) Hafsteinsdottir TB, Van Meeteren N, Gryphonck M. Is Bobath-based therapy used by physiotherapists in the care of stroke patients? *Adv Physiother* 2005 12;7(4):176-182.
- (35) Oostendorp RAB, Pluimers DJ, Nijhuis-Van der Sanden MWG, Wensing MJP. Fysiotherapeutische verslaglegging: de Achilleshiel voor Evidence-based Practice (EBP)? [Physiotherapy patient documentation: the Achilles heel of evidence-based practice?]. *Ned Tijdschr Fysiother* 2006;116(3):56-51.
- (36) G. M. J. Rutten, S. Degen, R. A. B. Oostendorp. Project Kwaliteitsindicatoren Fysiotherapie: Aspecifieke nekpijn en DTF. [Quality Indicators for Physiotherapy Project (ProKwaF): Nonspecific neck pain and direct access to physical therapy]. Nijmegen: Scientific Institute for Quality of Healthcare (IQ healthcare); 2007.
- (37) Turner PA, Harby-Owren H, Shackelford F, So A, Fosse T, Whitfield TWA. Audits of physiotherapy practice. *Physiother Theory Pract* 1999;15(4):261-274.
- (38) Engers AJ, Schers H, Oostendorp RAB, Wensing M, Koes BW, Grol R. Het handelen van Nederlandse fysiotherapeuten bij lage-rugpijn: een beschrijvend onderzoek van de

dagelijkse praktijk.[Management of low back pain in the Netherlands]. Ned Tijdschr Fysiother 2001 12;111(6):142-151.

(39) Li LC, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. Phys Ther 2001 04;81(4):1018-1028.

(40) Gracey JH, McDonough SM, Baxter GD. Physiotherapy management of low back pain: a survey of current practice in northern Ireland. Spine (Phila Pa 1976) 2002 Feb 15;27(4):406-411.

(41) Reid D, Larmer P, Robb G, Hing W, McNair P. Use of a vignette to investigate the physiotherapy treatment of an acute episode of low back pain: report of a survey of New Zealand physiotherapists. NZ J Physiother 2002 07;30(2):26-32.

(42) Armstrong MP, McDonough S, Baxter GD. Clinical guidelines versus clinical practice in the management of low back pain. Int J Clin Pract 2003 Jan-Feb;57(1):9-13.

(43) Hamm L, Mikkelsen B, Kuhr J, Stovring H, Munck A, Kragstrup J. Danish physiotherapists' management of low back pain. Adv Physiother 2003;5(3):109-113.

(44) Van der Linden SC, Lucassen M, Senders MM, Wiche LAJ, De Louw DCP. KNGF-richtlijn Lage-rugpijn: theorie in praktijk gebracht? [Dutch guideline for low back pain: theory brought into practice?]. FysioPraxis 2004;Februari:36-39.

(45) Mikhail C, Korner-Bitensky N, Rossignol M, Dumas JP. Physical therapists' use of interventions with high evidence of effectiveness in the management of a hypothetical typical patient with acute low back pain. Phys Ther 2005 Nov;85(11):1151-1167.

(46) Swinkels IC, Van den Bosch W, Dekker J, Wimmers RH. Physiotherapy management of low back pain: does practice match the Dutch guidelines? Aust J Physiother 2005;51(1):35-41.

(47) Flynn TW, Wainner RS, Fritz JM. Spinal manipulation in physical therapist professional degree education: A model for teaching and integration into clinical practice Fontys Eindhoven zie ook in hetzelfde issue Patient Outcomes and Clinical Performance: Parallel Paths or Inextricable Links? J Orthop Sports Phys Ther 2006 08;36(8):577-587.

(48) Fritz JM, Cleland JA, Brennan GP. Does adherence to the guideline recommendation for active treatments improve the quality of care for patients with acute low back pain delivered by physical therapists? Med Care 2007 Oct;45(10):973-980.

(49) Leerar PJ, Boissonnault W, Domholdt E, Roddey T. Documentation of red flags by physical therapists for patients with low back pain. J Man Manip Ther 2007;15(1):42-49.

(50) Bishop A, Foster NE, Thomas E, Hay EM. How does the self-reported clinical management of patients with low back pain relate to the attitudes and beliefs of health care practitioners? A survey of UK general practitioners and physiotherapists. Pain 2008 03;135(1-2):187-195.

- (51) Casserley-Feeney SN, Bury G, Daly L, Hurley DA. Physiotherapy for low back pain: differences between public and private healthcare sectors in Ireland--a retrospective survey. *Man Ther* 2008 10;13(5):441-449.
- (52) Ladeira CE, VanDenBerg GL. Physical therapist's nationality affects adherence to low back pain guidelines Fontys Eindhoven UU. *Journal of Orthopaedic & Sports Physical Therapy* 2009;39(1):A99-100.
- (53) Liddle SD, David BG, Gracey JH. Physiotherapists' use of advice and exercise for the management of chronic low back pain: a national survey. *Man Ther* 2009 04;14(2):189-196.
- (54) Rutten GMJ, Kremers SPJ, Rutten STJ, Harting J. A theory-based cross-sectional survey demonstrated the important role of awareness in guideline implementation. *J Clin Epidemiol* 2009 02;62(2):167-176.
- (55) Graham M. Frequency and duration of handwashing in an intensive care unit. *Am J Infect Control* 1990 Apr;18(2):77-81.
- (56) Pettinger A, Nettleman MD. Epidemiology of isolation precautions. *Infect Control Hosp Epidemiol* 1991 May;12(5):303-307.
- (57) Alexander KM, Sirotnak N. Self-reported hand-washing practices of Idaho State University physical therapy graduates. *J Phys Ther Educ* 1997 1997;11(1):3-9.
- (58) Bischoff WE, Reynolds TM, Sessler CN, Edmond MB, Wenzel RP. Handwashing compliance by health care workers: The impact of introducing an accessible, alcohol-based hand antiseptic. *Arch Intern Med* 2000 Apr 10;160(7):1017-1021.
- (59) Devine NL, Alexander KM, Peterson TS. Self-reported handwashing behaviors of practicing physical therapists. *Acute Care Perspectives* 2002 06;11(2):14-19.
- (60) Harbarth S, Pittet D, Grady L, Zawacki A, Potter-Bynoe G, Samore MH, et al. Interventional study to evaluate the impact of an alcohol-based hand gel in improving hand hygiene compliance. *Pediatr Infect Dis J* 2002 Jun;21(6):489-495.
- (61) Randle J, Clarke M, Storr J. Hand hygiene compliance in healthcare workers. *J Hosp Infect* 2006 Nov;64(3):205-209.
- (62) Santana SL, Furtado GHC, Coutinho AP, Medeiros EAS. Assessment of healthcare professionals' adherence to hand hygiene after alcohol-based hand rub introduction at an intensive care unit in Sao Paulo, Brazil. *Infect Control Hosp Epidemiol* 2007;28(3):365-367.
- (63) Creedon SA. Hand hygiene compliance: exploring variations in practice between hospitals. *Nurs Times* 2008 12/09;104(49):32-35.
- (64) Hammond R, Lennon S, Walker MF, Hoffman A, Irwin P, Lowe D. Changing occupational therapy and physiotherapy practice through guidelines and audit in the United Kingdom. *Clin Rehabil* 2005 06;19(4):365-371.

- (65) Dominique Cadilhac, Graeme Hankey, Dawn Harris, Susan Hillier, Monique Kilkenny, Erin Lalor. National Stroke Audit Clinical Report Acute Services. Melbourne: National Stroke Foundation; 2007.
- (66) Tyson SF, Selley AB. The effect of perceived adherence to the Bobath concept on physiotherapists' choice of intervention used to treat postural control after stroke. *Disabil Rehabil* 2007;29(5):395-401.
- (67) Van Peppen RP, Maissan FJ, Van Genderen FR, Van Dolder R, Van Meeteren NL. Outcome measures in physiotherapy management of patients with stroke: a survey into self-reported use, and barriers to and facilitators for use. *Physiother Res Int* 2008 12;13(4):255-270.
- (68) Luker J, Grimmer-Somers K. Factors influencing acute stroke guideline compliance: a peek inside the 'black box' for allied health staff. *J Eval Clin Pract* 2009 04;15(2):383-389.
- (69) Jansen MJ, Hendriks HJM, De Bie RA, Dekker J, Oostendorp RAB. Handelen volgens de KNGF-richtlijn 'Artrose heup-knie': een prospectieve cohortstudie. [Treatment according to the Dutch guideline on osteoarthritis of the hip and the knee: a prospective cohort study]. *Ned Tijdschr Fysiother* 2005;3(115):68-73.
- (70) Holden MA, Nicholls EE, Hay EM, Foster NE. Physical therapists' use of therapeutic exercise for patients with clinical knee osteoarthritis in the United Kingdom: In line with current recommendations? *Phys Ther* 2008;88(10):1109-1121.
- (71) Jamtvedt G, Dahm KT, Holm I, Flottorp S. Measuring physiotherapy performance in patients with osteoarthritis of the knee: a prospective study. *BMC Health Serv Res* 2008 Jul 8;8:145.
- (72) Walsh NE, Hurley MV. Evidence based guidelines and current practice for physiotherapy management of knee osteoarthritis. *Musculoskeletal Care* 2009 03;7(1):45-56.
- (73) Martin CJ, McCallum HM, Heaton B. An evaluation of radiofrequency exposure from therapeutic diathermy equipment in the light of current recommendations. *Clin Phys Physiol Meas* 1990 Feb;11(1):53-63.
- (74) Shields N, Gormley J, O'Hare N. Short-wave diathermy: current clinical and safety practices. *Physiother Res Int* 2002 12;7(4):191-202.
- (75) Shields N, O'Hare N, Boyle G, Gormley J. Development and application of a quality control procedure for short-wave diathermy units. *Med Biol Eng Comput* 2003 Jan;41(1):62-68.
- (76) Shah SG, Farrow A. Investigation of practices and procedures in the use of therapeutic diathermy: a study from the physiotherapists' health and safety perspective. *Physiother Res Int* 2007 /;12(4):228-241.
- (77) Roebroek ME, Dekker J, Oostendorp R, Bosveld W. Physiotherapy for patients with lateral ankle sprains: a prospective survey of practice patterns in Dutch primary health care. *Physiotherapy* 1998 09;84(9):421-432.

- (78) Larmer P, Robb G, Hing W, Reid D, McNair P. Use of a vignette to investigate the physiotherapy treatment of an acute ankle sprain: report of a survey of New Zealand physiotherapists. *NZ J Physiother* 2002 07;30(2):36-43.
- (79) Leemrijse CJ, Plas GM, Hofhuis H, Van den Ende CHM. Compliance with the guidelines for acute ankle sprain for physiotherapists is moderate in the Netherlands: an observational study. *Aust J Physiother* 2006;52(4):293-299.
- (80) Bennett S, Pescatello LS, Considine W, Cahill K. A Survey of New England Cardiac Rehabilitation Supervisors and Educational Degree Attainment. *Journal of Cardiopulmonary Rehabilitation and Prevention* 1994;14(5).
- (81) Bennett SB, Pescatello LS. A regional comparison of cardiac rehabilitation personnel: adherence to the 1995 American Association of Cardiovascular and Pulmonary Rehabilitation guidelines by staff position. *J Cardiopulm Rehabil* 1997 03;17(2):92-102.
- (82) Neeleman-Van der Steen CWM, Hendriks HJM, Bertram RJJ, Graus J, Van Herwaarden F, Jongert MWA, et al. Belangrijkste aanbevelingen KNGF-richtlijn Hartrevalidatie worden nageleefd. Resultaten van een prospectieve cohortstudie. [Compliance with the main recommendations of the Dutch cardiac rehabilitation guideline: results of a prospective cohort study]. *Ned Tijdschr Fysiother* 2008;118(1):2-11.
- (83) Magarey ME, Rebbeck T, Coughlan B, Grimmer K, Rivett DA, Refshauge K. Pre-manipulative testing of the cervical spine review, revision and new clinical guidelines. *Man Ther* 2004 05;9(2):95-108.
- (84) Abrams D, Davidson M, Harrick J, Harcourt P, Zylinski M, Clancy J. Monitoring the change: current trends in outcome measure usage in physiotherapy. *Man Ther* 2006 Feb;11(1):46-53.
- (85) Geraets JJ, Goossens ME, Van Haastregt JCM, De Groot I, De Bruijn CP, De Bie RA, et al. Implications of process evaluation for clinical effectiveness and clinical practice in a trial on chronic shoulder complaints. *Patient Educ Couns* 2006 04;61(1):117-125.
- (86) Gough S, Doherty J. Emergency on-call duty preparation and education for newly qualified physiotherapists: a national survey. *Physiotherapy* 2007;93(1):37-44.
- (87) Jones A, Hutchinson RC, Oh TE. Chest physiotherapy practice in intensive care units in Australia, the UK and Hong Kong. *Physiother Theory Pract* 1992 03;8(1):39-47.
- (88) Guillodo Y, Saraux A. Treatment of muscle trauma in sportspeople (from injury on the field to resumption of the sport). *Ann Phys Rehabil Med* 2009 04;52(3):246-255.
- (89) Meijer EM, Sluiter JK, Heyma A, Sadiraj K, Frings-Dresen MH. Cost-effectiveness of multidisciplinary treatment in sick-listed patients with upper extremity musculoskeletal disorders: a randomized, controlled trial with one-year follow-up. *Int Arch Occup Environ Health* 2006 09;79(8):654-664.

- (90) Nijkraak MJ, Keus SH, Ewalds H, Overeem S, Braspenning JC, Oostendorp RA, et al. Quality indicators for physiotherapy in Parkinson's disease. *Eur J Phys Rehabil Med* 2009 06;45(2):239-245.
- (91) Stokes EK, Bourke A, Monaghan F, Scully C. Hip protectors--a survey of practice in Ireland. *Ir Med J* 2005 01;98(1):21-23.
- (92) Gumery Y, Edenborough F, Stableforth D, Strachan A. Physiotherapy and nebuliser use in a Birmingham adult cystic fibrosis unit. *Physiotherapy* 1998;84(3):127-132.
- (93) Armantrout EA, Hummel-Berry K, Underwood F, Nelson C. Physical therapist compliance with electromyography guidelines. *J Neurol Phys Ther* 2008 12;32(4):177-185.
- (94) Rabey M, Morgans S, Barrett C. Orthopaedic physiotherapy practitioners: surgical and radiological referral rates. *Clinical Governance: An International Journal* 2009 03;14(1):15-19.
- (95) Schonstein E, Kenny DT, Maher CG. WorkCover's physiotherapy forms: purpose beyond paperwork? *Aust J Physiother* 2002;48(3):221-225.
- (96) Baker SM, Marshak HH, Rice GT, Zimmerman GJ. Patient participation in physical therapy goal setting. *Phys Ther* 2001 05;81(5):1118-1126.
- (97) Saliga S, Adamowicz C, Logue A, Smith K. Physical therapists' knowledge of physical elder abuse -- signs, symptoms, laws, and facility protocols auteur zoeken. *Journal of Geriatric Physical Therapy* 2004 03;27(1):3-10.
- (98) Hrachovy J, Clopton N, Baggett K, Garber T, Cantwell J, Schreiber J. Use of the Blue MACS: Acceptance by clinical instructors and self-reports of adherence. *Phys Ther* 2000;80(7):652-661.
- (99) Kwakkel G, Wagenaar RC, Twisk JW, Lankhorst GJ, Koetsier JC. Intensity of leg and arm training after primary middle-cerebral-artery stroke: a randomised trial. *Lancet* 1999 Jul 17;354(9174):191-196.
- (100) Berwick DM. Continuous improvement as an ideal in health care. *N Engl J Med* 1989 Jan 5;320(1):53-56.
- (101) Grol R, Wensing M. Implementatie, effectieve verandering in de patiëntenzorg [Implementation, Effective Changes in Patient Care]. 2nd ed. Maarssen: Elsevier gezondheidszorg; 2001.
- (102) Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. *BMJ* 2003 Apr 12;326(7393):816-819.
- (103) McGlynn EA, Asch SM. Developing a clinical performance measure. *Am J Prev Med* 1998 Apr;14(3 Suppl):14-21.

- (104) LaClair BJ, Reker DM, Duncan PW, Horner RD, Hoenig H. Stroke care: a method for measuring compliance with AHCPR guidelines. *Am J Phys Med Rehabil* 2001 Mar;80(3):235-242.
- (105) Stuurgroep Indicatorontwikkeling. Ontwikkeling van Indicatoren op Basis van Evidence-Based Richtlijnen [Development of Indicators Based on Evidence-Based Guidelines]. Utrecht: Kwaliteitsinstituut voor de Gezondheidszorg CBO; 2002.
- (106) Landau SI, Bogus RJ.
The New International Webster's Student Dictionary of the English Language: International Encyclopedic Edition. : Trident Press International; 1996.
- (107) Merriam-Webster's Learner's Dictionary. Available at: <http://www.learnersdictionary.com/search/adhere>. Accessed 6/2/2011, 2011.
- (108) adhere, v. : Oxford English Dictionary. Available at: <http://www.oed.com.proxy.library.uu.nl/view/Entry/2327>. Accessed 6/2/2011, 2011.
- (109) Braspenning J, Hermens R, Wollersheim H, Grol R. Meten van (verandering in) de zorg: de rol van indicatoren [Measuring of (Change in) in Care: the Role of Indicators]. In: Grol R, Wensing M, editors. Implementatie, effectieve verbetering van de patiëntenzorg [Implementation, Effective Enhancement of Patient Care]. 3rd Rev ed. Maarssen: Elsevier gezondheidszorg; 2006. p. 153-179.
- (110) Casparie AF, Hommes H. Indicatoren voor integrale kwaliteitszorg in ziekenhuizen [Indicators for integrated quality of care in hospitals]. Deventer: Kluwer; 1997.
- (111) Lindsay P, Schull M, Bronskill S, Anderson G. The development of indicators to measure the quality of clinical care in emergency departments following a modified-delphi approach. *Acad Emerg Med* 2002 Nov;9(11):1131-1139.
- (112) Brook RH. The Rand/UCLA Appropriateness Method. In: McCormick KA, Moore SR, Siegel RA, editors. Clinical Practice Guideline Development: Methodologic Perspectives Rockville, MD: Public Health Service, US Department of Health and Human Services. AHCPR Publication No. 95- 0009; 1994. p. 59-65.
- (113) Holloway RG, Vickrey BG, Benesch C, Hinchey JA, Bieber J, National Expert Stroke Panel. Development of performance measures for acute ischemic stroke. *Stroke* 2001 Sep;32(9):2058-2074.
- (114) Baker R, Fraser RC. Development of review criteria: linking guidelines and assessment of quality. *BMJ* 1995 Aug 5;311(7001):370-373.
- (115) Nelen W, Kremer J, Hermens R, Braat D, Grol R. Development of quality indicators: A first step in the implementation of Dutch infertility clinical guidelines. *Tijdschr Fertil* 2003 2003/12;17(4):90-95.
- (116) J. S. De Koning, A. W. Smulders, N. S. Klazinga. Appraisal of Indicators through Research and Evaluation (Aire) Instrument (Dutch). Amsterdam: Academisch Medisch Centrum; 2007.

