Forschende Komplementärmedizin
Wissenschaft • Praxis • Perspektiven

Review Article · Übersichtsarbeit

Forsch Komplementärmed 2008;15:40–46
DOI: 10.1159/000112820

Published online: January 28, 2008

A Review of the Literature in Applied and Specialised Kinesiology

Sue Hall George Lewith Sarah Brien Paul Little

Complementary Medicine Research Unit, School of Medicine, University of Southampton, UK

Key Words
Review · Applied Kinesiology · Muscle test

Summary
Introduction: Kinesiology is a diagnostic, therapeutic complementary therapy utilising subtle change in manual muscle testing results to evaluate the body's energetic balance and select healing modalities. Anecdotal evidence suggests kinesiology is helpful, therefore we wished to critically review the literature. **Aims:** (1) To ascertain if diagnostic accuracy including inter-examiner reliability has been established. (2) To review whether there is evidence for its therapeutic effectiveness. (3) To critically assess the quality of relevant studies. **Methods:** Electronic databases were searched. Diagnostic accuracy studies were analysed and scored for methodological quality and quality of reporting using the quality assessment tool for studies of diagnostic accuracy included in systematic reviews (QUADAS) and the Standards for Reporting of Diagnostic Studies (STARD). Clinical studies were analysed for methodological quality using the JADAD scale and for quality of reporting using the Consolidated Standards of Reporting Trials (CONSORT). **Results:** 22 original relevant studies were identified. Their methodology was poor. Items reported on QUADAS scored 1–11 out of a possible 14, STARD scores were between 6–13 out of 25, JADAD scores were all 0 out of 5 and CONSORT 4–6 out of 22. Consequently, we were unable to answer any of our research questions. **Conclusion:** There is insufficient evidence for diagnostic accuracy within kinesiology, the validity of muscle response and the effectiveness of kinesiology for any condition. The standards of reporting were low. We recommend a pragmatic study of the effectiveness of kinesiology as the most appropriate initial step to determine whether kinesiology has any clinical value.

Schlüsselwörter
Literaturstudie · Angewandte Kinesiologie · Muskeltest

Zusammenfassung
Einleitung: Die Kinesiologie versteht sich als komplementärmedizinisches Diagnose- und Therapieverfahren. Dabei dienen kleinste Veränderungen im manuellen Muskeltest der Beurteilung des energetischen Gleichgewichts des Körpers und der Auswahl anzuwendender Heilmethoden. Da Einzelfälle die Kinesiologie als hilfreich beschreiben, wollten wir die vorliegende Literatur kritisch untersuchen. **Ziele:** (1) Festzustellen, ob diagnostische Genauigkeit und Reliabilität zwischen verschiedenen Untersuchern besteht. (2) Zu prüfen, ob es Evidenz für die therapeutische Wirksamkeit gibt. (3) Kritische Beurteilung der Qualität der vorliegenden Studien. **Methoden:** Elektronische Datenbanken wurden durchsucht. Diagnostische Genauigkeitsstudien wurden mit Hilfe von QUADAS (Instrument zur Qualitätsabschätzung von Studien zur diagnostischen Genauigkeit) und STARD (Standards for Reporting of Diagnostic Studies) beurteilt. Die methodische Qualität der klinischen Studien wurde mit Hilfe des JADAD beurteilt, die Qualität der Studienberichte mit dem CONSORT-Statement (Consolidated Standards of Reporting Trials). **Ergebnisse:** 22 relevante Studien konnten identifiziert werden. Die Methodik war bei allen Studien minderwertig. Die erreichte Qualität im QUADAS lag bei 1–11 von möglichen 14 Punkten, im STARD bei 6–13 von maximal 25, im JADAD bei 0 von maximal 5 Punkten. Die Qualitätsbeurteilung anhand des CONSORT-Statement ergab 6–22 von maximal 22 Punkten. Demzufolge war es uns nicht möglich, unsere Forschungsfragen zu beantworten. **Schlussfolgerung:** Beweise für die diagnostische Genauigkeit, Aussagekraft der Muskelreaktion und therapeutische Wirksamkeit der Kinesiologie sind unzureichend. Die Qualität der Studienberichte ist mangelhaft. Es sollte zunächst eine pragmatische Studie zur Wirksamkeit der Kinesiologie durchgeführt werden, um einen möglichen klinischen Nutzen festzustellen.

KARGER

Fax +49 761 4 52 07 14
E-mail Information@Karger.de
www.karger.com

Accessible online at:
www.karger.com/fok

Sue Hall MA
The British Kinesiology Centre, Alexandra House
1–3 Waverley Lane, Farnham
Surrey, GU9 8BB, United Kingdom
Tel. +44 1252 722 246, Fax 2380 701125
E-mail s.s.Hall@soton.ac.uk

Introduction

Applied kinesiology (AK), initially developed by George Goodheart in the 1960's, is a chiropractic speciality used in conjunction with chiropractic technique and combined with a standard physical examination, x-rays, history taking and laboratory findings. AK is both a diagnostic and therapeutic system; it utilises manual muscle testing to assess change in neuromuscular function in response to physical, chemical or mental stimuli. The history, development and detailed diagnostic and therapeutic processes of applied kinesiology are described elsewhere [1, 2]. Subsequently, in the 1970's John Thie developed a simple offshoot of AK for lay people to use at home called Touch for Health Kinesiology (TFH) [3]. Numerous expanded variations of this simplified method were developed, some of which utilise a light muscle test as a yes/no answer system (strong = yes, weak = no) and derive their therapeutic interventions from energetic healing theories. These systems became known as the 'specialised or energy kinesiologies' [4], some of which are considered professions in their own right. All kinesiology systems derive from AK and generally, all types of kinesiology including AK are known to the public as 'kinesiology'. Each branch is said to differ slightly in its approach although the basic premises are the same for all kinesiology and there are very many similarities in technique between the different kinesiologies. We developed generic questions for this literature review that can reasonably be asked of all these kinesiology approaches and therefore all types of kinesiology are included.

A review of the literature published by the International College of Applied Kinesiology (ICAK) between 1981–1987 [5] concluded that the methodological quality of the studies was poor and that there was no justification for the conclusions drawn from the reported findings. A more recent literature review published after we began this review concluded that there is evidence to support the validity of manual muscle testing; this has been interpreted positively by the kinesiology community. However, the conclusions may have been misunderstood as the studies identified were not critically appraised for internal and external validity [6]. Further kinesiology studies have been published, so we felt that it was timely to update the 1990 review. As a consequence, we identified papers to address the following aims and specific research questions: (1) To ascertain if diagnostic accuracy for any type of kinesiology has been established. (2) To ascertain if inter-examiner reliability for any type of kinesiology has been established. (3) To ascertain if muscles respond to stimuli as suggested by the underlying principles of kinesiology. (4) To review whether there is any evidence for the therapeutic effectiveness of kinesiology. Additional Aim: To examine the quality of reporting of the studies retrieved.

Methods

We undertook a review of kinesiology within the context of complementary and alternative medicine (CAM) including all branches of kinesiology, as we suspected that there would only be a small amount of literature. From clinical experience (SH), we defined kinesiology as a method using subtle change in manual muscle testing results to assess the energetic balance of the body and to subsequently select individualised healing modalities.

Search

A search was conducted of Medline, Embase, CINHAL and AMED without restriction for date or language. The key words 'kinesiology', 'applied kinesiology', 'specialised kinesiology' and 'manual muscle testing' were used to identify appropriate papers for inclusion. Titles, key words and abstracts were read and a full copy of each paper that appeared to be relevant was retrieved and translated, if necessary, and citation tracked for further studies. Where we were unsure from the abstract whether to include or exclude the study, the full paper was obtained for clarification. We identified papers within the grey literature by contacting the kinesiology associations, by contacting kinesiology practitioners (SH), from kinesiology websites and from hand searching kinesiology conference proceedings. Where the study appeared relevant, individual researchers were contacted for further details and full papers were obtained. SH carried out the literature search, took the final decisions on the inclusion or exclusion of studies where necessary and led the analysis in collaboration and with advice from SB, GL and PL.

Inclusion and Exclusion Criteria

We included any clinical trial of any type of kinesiology for any outcome on patients or volunteers and any studies fulfilling the criteria identified by our aims. We excluded the mechanical measurement of muscle strength as not representative of our definition of kinesiology practice. We excluded muscle strength tests as in orthopaedic testing as kinesiology testing measures the ability of the nervous system to adapt to the light pressure of the test [2, 7] rather than the power a muscle produces.

Criteria for Quality Assessment

Clinical trials reporting treatment effectiveness or efficacy were evaluated for quality using the JADAD scale [8], and quality of report using the CONSORT statement [9]. The JADAD scale is a widely used, validated tool to assess the rigour of clinical trials included in systematic reviews [10]. It consists of 5 questions scoring either 0 or 1. A score of 0–2 is poor, 3–4 is good and 5 is banded as excellent [11]. The CONSORT statement consists of a 22-item checklist and flow diagram offering guidance to peer reviewers assessing the quality of reporting of such trials. A scoring system was not available, so we noted the number of CONSORT items reported out of 22 with the assumption that a lower number of reported items represented a lower quality of report.

QUADAS [12] was used to assess the quality of diagnostic accuracy studies, and the STARD criteria [13] to assess the quality of report. QUADAS consists of a validated list of 14 questions designed to assess the quality of studies of diagnostic accuracy included in systematic reviews [14]; the questions are answered yes, no or unclear. There was no scoring system available for QUADAS, therefore we noted the items reported with the assumption that a lower number of reported items would represent a lower quality study. The STARD initiative consists of a 25-item checklist and flow diagram designed to improve the quality of reporting for studies of diagnostic accuracy. Similarly, no scoring system was available, therefore we noted the items reported in the same way. Summing the number of STARD items to assess quality has been used previously [15] although to date this method has not been validated. We assumed that studies reporting a low number of items on either quality assessment tool indicated a low quality of report.

Table 1. Reported characteristics, methodological and reporting quality items noted of included studies (grouped by type of study, alphabetical order)

Studies grouped by type (1st author, year)	Participants	Examiners	Procedure	Reported results, (reported statistics)	Methodological quality items (better rated studies have a higher score)	Reporting items (better rated studies have a higher score)
Diagnostic accuracy					*QUADAS items reported out of 14*	*STARD items reported out of* 25
Conable et al 2006	30 chiropractic students/spouses	2 AK practitioners > 20 yrs experience	AK muscle test results vs lab tests for adverse reactions to foods	no significant correlation between AK and lab tests (kappa < 0.22 for 5 foods)	10	11
Jacobs et al 1984	65 patients and volunteers, some with symptomology	1 AK practitioner, 1 chiropractor	AK muscle test vs lab test vs chiropractic clinic protocol	correlations were significant but below optimum (all Spearman's correlations between different ratings ≤ 0.47)	1	10
Kenney et al 1988	11 volunteers	1 AK chiropractor, 2 people said to be experienced in AK	AK muscle test vs lab test for nutritional deficiency	high frequencies of false negative and false positive compared to lab tests (no statistics for diagnostic accuracy tests)	7	7
Ludtke et al 2001	7 patients with history of anaphylaxis after insect stings	4 health kinesiologists of varying experience	Double blind muscle test. Patient held glass vials of wasp venom solution vs salt solution	no inter-examiner reliability. Kinesiology not more useful than random guessing (kappa = 0.03 (CI 0.02–0.07). Sensitivity/specificity estimates 40% and 60%)	11	13
Pothmann et al 2001	315 < 18 yrs with chronic diseases; eczema, headache, asthma etc	4 testers	AK muscle tests vs lab tests for adverse reactions to foods	no inter-examiner reliability (kappa = –0.01) no significant agreement with lab tests (73% sensitivity, 43.2% specificity)	11	12
Rybeck et al 1980	73 volunteers	1 examiner	Muscle test with sugar cube in mouth vs no sugar. Similarly with a mechanical test.	manual testing significantly different between sugar and control. (Wilcoxon rank sum p = 0.0062) Mechanical test not significant	9	6
Schmitt et al 1999	17 volunteers	not specified	AK test vs lab tests for adverse reactions to foods	90.5% of positive lab results correlated with AK tests. (Cochrons Q p < 0.0001)	6	7
Mean reported items					7.8	9.4
Inter-examiner reliability					*Basic methodological criteria reported out of 6*	*Quality of report based on same 6 criteria*
Chung et al 2005	10 dental students	2 examiners	AK muscle test to determine arm strength related to jaw position	significant correlation on first test (p ≤ 0.023), no correlation on repeated test (p ≥ 0.09)	3	3
Hass et al 1993	68 volunteers	2 AK testers	AK muscle test change when force to spinous process applied	examiner concordance attributable to chance	5	5
Jacobs et al 1981	100 AK seminar attendees 10 volunteer students	2 AK testers	AK test with sugar solution vs distilled water AK test with sugar solution vs lump sugar AK muscle test with old vs new oils	no significant difference (^{2} p > 0.05) comparison of sugar solution vs lump sugar – significant difference (^{2} 0 < 0.05) comparison of old vs new oil: no significant difference (^{2} p > 0.03)	1	1

Lawson et al 1997	32 volunteers – 1st test group 53 volunteers – 2nd test group	3 AK testers > 10 yrs experience	muscle testing of 4 muscle groups = strong or weak result	good examiner concordance for 2 muscle tests. (kappa > 0.70 $p < 0.001$ for piriformis) no significant difference for other muscles	3	3
Peterson 1996	27 phobic volunteers	2 chiropractors > 10 yrs experience of muscle testing	AK test in relation to a word on a card, e.g. snake	inter-examiner reliability attributable to chance (kappa = –0.19 subjects combined)	0	0
Scopp 1979	10 subjects	6 trained testers	muscle test = weak or strong muscle test change in relation to nutrient or placebo muscle test vs Philpot-type fast	kinesiology testing may be statistically reliable and valid (Pearson's r = 0.91) ($p < 0.05$) (r = 0.81)	0	0
Staehle et al 2005	40 participants	2 AK dentists	AK muscle test for tolerance or non tolerance to dental materials	accuracy of AK testing – not significant ($p = 0.97$)	3	3
Mean reported items					2.1	2.1
Muscle response						
Friedman et al 1981	19 college students 20 college students 16 college students	4 examiners	AK test on dental occlusal position AK test for kinesiology technique AK test to sugar and Vitamin E.	3 students seem stronger 5 students tested weaker some weaker, some stronger (none reported)	1	1
Garrow 1988	20 patients	1 tester	patients with reaction to food test on open testing were blind tested for that food	kinesiology response not reproducible on blind testing (none reported)	2	2
Gast et al 1994	29 students	1 examiner	blind muscle test with and without a magnet on acu pt Spleen 5	significant difference between magnet in triceps test ($p = 0.0184$)	1	1
Monti et al 1999	89 subjects	1 tester	time by force trials of muscle testing in response to congruent and incongruent statements	congruent statements associated with significantly higher scores of time and force vs incongruent ones ($p < 0.001$)	3	3
Triano 1982	50 patients per examiner	unclear 4 or 5 examiners	tests to strengthen the latissimus dorsi by topical or sublingual exposure to food supplements	in blind test no consistent relationship between specific nutrient making a muscle stronger (chi square analysis failed to reveal any preponderant response as claimed by AK)	0	0
Mean reported items					0.9	0.9
Effectiveness					*Jadad items reported out of 5*	*CONSORT items reported out of 22*
Dawson 1999	1 clients	15 (3 in 1 consultants)	1–22 consultations of kinesiology measuring perceived stress before and after sessions	significant reduction in stress and ability to cope as measured by the perceived stress scale (PSS) (reduction in mean score $p < 0.0005$)	0	4
Gregory et al 2001	88 patients with mastalgia	2 kinesiologists	self-assessed pain levels before and after a kinesiology technique	significant reduction of pain after first treatment (t-test $p < 0.00001$)	0	4
Webb 1993	26 subjects with recurring dreams	1 experimenter	description of dream with or without (control) interpretation using kinesiology	mean dream frequency reduced in treatment group (analysis of variance $p < 0.006$)	0	6
Mean reported items					0	4.6

No validated scoring system for methodological quality or quality of reporting was available for inter-examiner reliability studies. We assessed the internal and external validity of these studies by examining for the following 6 basic methodological criteria: randomisation, inclusion criteria, blinding of examiners to each others results, representative sample of examiners, appropriate statistics, and sample size.

Results

Searches

Searches of Medline generated 26 potentially relevant kinesiology studies and AHMED retrieved 3 further studies. EMBASE and Cinahl produced duplicates and citation tracking did not yield further relevant studies. On retrieval of full articles we subsequently excluded 10 studies; 1 incorporated other clinical therapeutics [16], 8 did not meet our inclusion criteria [17–24] and 1 tested a premise not included in kinesiology practice [25]. Practitioner contact produced 2 studies [26, 27]. Kinesiology websites produced 4 further potentially relevant studies [28–31] of which 1 privately published study [29] was retrieved directly from the author and the other 3 [28, 30, 31] were excluded as not meeting the inclusion criteria. A total of 22 studies matched our inclusion and exclusion criteria and related to our aims (table 1). Of these papers, the majority (n = 18) cited AK as the type of kinesiology studied, however, 3 studies [26, 32, 33] utilised branches of specialised kinesiology, and 1 study [34] did not identify the type of kinesiology used.

Diagnostic Accuracy Studies

7 diagnostic accuracy studies were identified [29, 32, 35–39]. 4 studies compared kinesiology muscle testing for food allergies with standard laboratory tests, e.g. IgG and RAST tests [29, 32, 37, 39]: 1 of these concluded that kinesiology testing was not better than guessing [32] and 1 concluded that AK does not appear to correlate with laboratory tests [40]. 1 study compared kinesiology muscle testing for nutrients with biochemical tests for nutrient status [36] but concluded that results could have been due to suggestion; 1 study compared kinesiology findings with chiropractic clinical observations and laboratory analysis for thyroid profile [35] and concluded that AK alone did not identify thyroid dysfunctionals; 1 study compared kinesiology muscle test results to mechanical muscle test results before and after the ingestion of sugar [38] suggesting positive results for manual testing. 3 studies [32, 35, 37] did not use a representative patient population or clearly reported selection criteria [36, 38, 39]. Blinding was inadequate in 3 studies [29, 38, 39]. The kinesiology test in [39] influenced the decision to perform the reference test which may have led to biased estimates of accuracy; this study was also uncontrolled, however, the authors concluded that AK may be of value in screening for food allergies. Only 3 studies [29, 32, 37] adequately reported statistical methods used to calculate measures of diagnostic accuracy such as sensitivity, specificity and positive predictive value. No study reported considerations for sample size and only 2 studies provided confidence intervals [32, 37]. Small sample size can lead to imprecise estimates of accuracy [41], and measures of precision are necessary to determine the range within which true values lie [42]. QUADAS items clearly reported and thus scoring 'yes' ranged from 1–11 out of 14 (table 1). Overall, study quality was not sufficient to ascertain that diagnostic accuracy for kinesiology has been established.

Inter-Examiner Reliability Studies

7 inter-examiner reliability studies were identified [43–49]. Studies compared the reliability of examiners to detect the presence of weak or strong muscles [44] concluding promising results, the sensitivity to dental materials or foods [46–49], to pressure on specific vertebrae [43] where it was concluded that the results were due to chance and to a phobic stimulus (e.g. the word snake) [45]. 4 studies did not clearly report randomisation [46–49] which could have lead to biased results, 5 studies did not mention inclusion criteria [44, 46–49] and in 3 studies the examiners were not adequately blinded [44, 46, 49]. All studies but 1 [46] reported the training and experience of the examiners and appropriate statistics but only 1 study mentioned confidence intervals [47]. One study described post hoc data manipulation, reclassifying the results and claiming valid findings [45], yet 1 study provided no methodological information at all whilst claiming reliability and validity for kinesiology testing [46]. Overall, the quality of studies was not sufficient for us to ascertain if inter-examiner reliability for kinesiology has been established (table 1).

Muscle Response to Stimuli

5 papers evaluating muscle response to stimuli studies were identified [27, 34, 50–52] of which 3 concluded positive results [27, 34, 51]. This group of papers tested the presence of weak or strong muscles in relation to foods [34, 50, 52], a kinesiology technique [50], magnets [27] and verbal statements [51]. Only 1 study described the examiner [51] and 1 study reported effective blinding and an appropriate randomisation procedure [34]. Inclusion criteria were reported in 3 studies [27, 50, 51], 1 uncontrolled study reported statistical data for only part of the experiment [51], but no study discussed confidence intervals or sample size. In 1 study it was unclear how many subjects were included in each experiment [50]. The studies were not of sufficient quality to determine if muscle responses observed were consistent with the underlying principles of kinesiology (table 1).

Therapeutic Effectiveness of Kinesiology

3 clinical outcomes studies were identified [26, 33, 53]. These studies assessed the effectiveness of kinesiology for mastalgia [53], stress [26] and recurring dreams [33]. None of the studies were randomised but 1 study described a control group [33]; there was no information about how subjects had been allocated to the control and it was debateable whether the control

was credible or subjects were blind to treatment. All the studies described significant results but it was unclear whether the result was related to the specific effects of treatment, as they were either uncontrolled [26, 53] or biased [33]. 2 studies used validated outcome measures but did not mention what constituted a clinically important difference [26, 53]. There was no mention of sample size in any study. The JADAD scale was used to score the methodological quality of these papers (table 1); all the studies scored 0 inferring that overall, the quality was poor [8]. Therefore, it was not possible to use these studies as pilots to evaluate power for further more definitive trials. We were thus unable to ascertain if there is any evidence for the therapeutic effectiveness of kinesiology.

Discussion and Conclusion

It is clear that efforts are needed to improve both methodological and reporting quality of studies in this field. The STARD criteria for diagnostic accuracy studies would encourage authors to use a more standardised and clear approach in reporting, and using QUADAS could improve study design. The use of CONSORT would help both in study design and quality of report of clinical outcomes studies. We also considered the following to be of importance:

The muscle test as a diagnostic tool is central to the practice of kinesiology but it is unclear whether it is valid or reliable. This cannot be evaluated rigorously unless a good model for validity is ensured. A recent literature review of the reliability and validity of manual muscle testing concluded that there is evidence to support the validity of manual muscle testing within AK, but the studies identified were not critically evaluated for internal and external validity and negative studies were not included: this conclusion is therefore premature [6]. We concur with the reviewers that a systematic review of the manual muscle testing studies would be pertinent at this time

Kinesiologists using differing systems may have fundamental differences in their use of, understanding and interpretation of muscle testing and this represents a challenge for ensuring model validity is good in studies in this field. A Delphi or similar consensus could establish whether there are real differences which are important for study design.

Patients presenting with ostensibly the same standard medical diagnosis would not necessarily be considered heterogeneous according to kinesiology assessment. This is a limitation scientifically and highlights the conflict between scientific rationale and the philosophy of kinesiology. Again, a Delphi or similar consensus could establish a basic population for testing, but we suggest that kinesiology treatment should be individualised and not standardised in any clinical trial.

It is inherently challenging to control for something that is as potentially subjective as muscle testing. Understanding the issues in sham and placebo controls and specific and non-specific effects in complex interventions such as kinesiology or acupuncture is challenging when there is poor understanding of mechanisms [54, 55]. This needs to be considered in detail when designing a clinical trial. Blinding to treatment is not possible for the practitioner although it may be possible to blind naive and experienced subjects to certain forms of treatment. This suggests that clinical trials should be pragmatic and single blind.

It is possible that we failed to identify some papers that were not within the peer reviewed literature, however, based on the quality of the studies we did examine, we believe that this was unlikely to have changed our results. Our definition of kinesiology may have led to our excluding potentially relevant studies; however, at present there is no model validity consensus within the field to have enabled us to be more precise. The type of muscle testing used by examiners appears to differ within and between branches of kinesiology [2, 3, 7, 56]. It is uncertain at this time whether this is important, but clearly some clarification is required. Lastly, the lack of available quality criteria tools for inter-examiner reliability studies or muscle response to stimuli studies necessitated us to define our own criteria.

Based on this review of the studies there is insufficient evidence to suggest that kinesiology (of any type) has any specific therapeutic effect for any condition; that inter-practitioner agreement in relation to a kinesiology diagnosis has been demonstrated or that the validity of muscle testing has been established. What the literature has demonstrated is that there are inherent difficulties in attempting research in this area but that steps to increase the rigour and generalisability of the experiments are needed. Anecdotal evidence suggests that kinesiology is a clinically helpful therapy. We suggest that kinesiology needs to be evaluated initially as a whole system with a controlled and rigorous but pragmatic approach. We clearly need to understand if this system is of any clinical value before we begin to evaluate its various components. We propose therefore that a pragmatic single blind randomised controlled study be conducted to assess the clinical and therapeutic effects of kinesiology.

Acknowledgements and Sources of Support

Dr Lewith's post is funded by a grant from the Rufford Maurice Laing Foundation. Dr Brien is supported by a Post-doctoral Personal Development Award funded by the Department of Health National Co-ordinating Centre for Research Capacity Development. Professor Little receives funding from the Higher Education Funding Council for England. Sue Hall – none.

Conflict of Interest

Sue Hall is an examiner for the International College of Professional Kinesiology Practice (ICPKP)

References

1 Green BN, Gin RH: George Goodheart DC and a history of applied kinesiology. J Manipulative Physiol Ther 1997;20:31–337.
2 Walther DC: Applied Kinesiology. Synopsis, ed 2. Pueblo, CO, Systems DC, 1988.
3 Thie JF: Touch for Health. Manna del Rey, CA, DeVorss, 1973.
4 Wolontis, Mac Pompeius: The Kinesiology Network. 2007, www.Kinesiology.net.
5 Klinkoski B, Leboeuf C: A review of the research papers published by the International College of Applied Kinesiology from 1981 to 1987. J Manipulative Physiol Ther 1990;13:190–194.
6 Cuthbert SC, Goodheart Jr GJ: On the reliability and validity of manual muscle testing: a literature review. Chiropr Osteopat 2007;15:4.
7 Dewe BAJ, Dewe JR: Five Element Fundamentals and Muscle Testing Basics for Energy Therapists. K-Power® Global Health and Fitness, 2007.
8 Jadad AR, Moore A, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, McQuay HJ: Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996; 17:1–12.
9 Moher D, Schulz KF, Altman DG: The Consort Statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001;357:1191–4.
10 van Tulder MW, Assendelft WJ, Koes BW, Bouter LM: Methodological guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. Spine 1997;22:2323–2330.
11 Natural Standards (ed): Explanation of columns in natural standard evidence table. www.natural-standard.com/explanation-columns.
12 Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PM: The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25.
13 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HCW, Lijmer JG: The STARD Statement for Reporting Studies of Diagnostic Accuracy: Explanation and Elaboration. Clin Chem 2003;49:7–18.
14 Whiting P, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J: Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol 2006;6:9.
15 Smidt N, Rutjes AWS van der Windt DAWM, Ostelo RWJG, Reitsma JB, Bossuyt PM, Bouter LM, de Vet HCW: Quality of Reporting of Diagnostic Accuracy Studies. Radiology 2005;235:347–353.
16 Moncayo R, Moncayo H, Ulmer H, Kainz H: New diagnostic and therapeutic approach to thyroid-associated orbitopathy base on applied kinesiology and homeopathic therapy. J Altern Complement Med 2004;10:643–650.
17 Caruso W, Leisman G: A force/displacement analysis of muscle testing. Percept Mot Skills 2000;91: 683–692.
18 Caruso W, Leisman G: The clinical utility of force/displacement analysis of muscle testing in applied kinesiology. Int J Neurosci 2001;106:147–157.
19 Hsieh C-Y, Phillips RB: Reliability of manual muscle testing with a computerized dynamometer. J Manipulative Physiol Ther 1990;13:72–82.
20 Leisman G, Shambaugh P, Ferentz AH: Somatosensory evoked potential changes during muscle testing. Int J Neurosci 1989;45:143–151.
21 Leisman G, Zenhausern R, Ferentz AH, Tefera T, Zemcov A: Electromyographic effects of fatigue and task repetition on the validity of estimates of strong and weak muscles in applied kinesiological muscle-testing procedures. Percept Mot Skills 1995; 80:963–977.
22 Perot C, Meldener R, Goubel F: Objective measurement of proprioceptive technique consequences on muscular maximal voluntary contraction during manual muscle testing. Aggressologie 1991;32:471–474.
23 Peter KB: The effects of spinal manipulation on the intensity of emotional arousal in phobic subjects exposed to threat stimulus: a randomised contolled, double-blind, clinical trial. J Manipulative Physiol Ther 1997;20:602–606.
24 Pollard H, Lakay B, Tucker F, Watson B, Bablis P: Interexaminer reliability of the deltoid and psoas muscle test. J Manipulative Physiol Ther 2005;28: 52–56.
25 Grossi JA: Effects of an applied kinesiology technique on quadriceps femoris muscle isometric strength. Phys Ther 1981;61:1011–1016.
26 Dawson S: Assisted healing by reduction of stress, using three in one concepts. An evaluation of clients' perceptions. 1999. Dissertation, Victoria University, Melbourne.
27 Gast AE, Bronks R: Angewandte Kinesiologie. Eine Doppelblindstudie zur Bewertung der Muskel-Meridian-Beziehung. Krankengymnastik 1994; 46:1020–1028.
28 Cammisa KM: Educational kinesiology with learning disabled children: an efficacy study. Percept Mot Skills 1994;78:105–106.
29 Conable K, Zhang J, Hambrick T: Comparison of applied kinesiology neuromuscular screening and laboratory indicators of adverse reactions to foods. The Collected Papers of the International College of Applied Kinesiology 2006;35–44.
30 Khalsa GK, Morris GSD, Sifft JM: Effect of educational kinesiology on static balance of learning disabled students. Percept Mot Skills 1988;67:51–54.
31 Sifft JM, Khalsa GCK: Effect of educational kinesiology upon simple response times and choice response times. Percept Mot Skills 1991;73: 1011–1015.
32 Ludtke R, Kunz B, Seeber N, Ring J: Test-retest-reliability and validity of the kinesiology muscle test. Complement Ther Med 2001;9:141–145.
33 Webb DE, Fagan J: The impact of dream interpretation using psychological kinesiology on the frequency of recurring dreams. Psychother Psychosom 1993;59:203–208.
34 Garrow JS: Kinesiology and food allergy. BMJ 1988;296:1573–1574.
35 Jacobs GE, Franks TL, Gilman PG: Diagnosis of thyroid dysfunction:applied kinesiology compared to clinical observations and laboratory tests. J Manipulative Physiol Ther 1984;7:99–104.
36 Kenney JJ, Clemens R, Forsythe KD: Applied kinesiology unreliable for assessing nutrient status. J Am Diet Assoc 1988;88:698–704.
37 Pothmann R, von Frankenberg S, Hoicke C, Weingarten H, Ludtke R: Evaluation der klinisch angewandten Kinesiologie bei Nahrungsmittelunverträglichkeit im Kindesalter. Forsch Komplementarmed Klass Naturheilkd 2001;8:336–344.
38 Rybeck CH, Swenson R: The effect of oral administration of refined sugar on muscle strength. J Manipulative Physiol Ther 1980;3:155–161.
39 Schmitt WH, Leisman G: Correlation of applied kinesiology muscle testing findings with serum immunogobulin levels for food allergies. Int J Neurosci 1999;96:237–244.
40 Conable K, Zhang J, Hambrick T: Comparison of applied kinesiology neuromuscular screening and laboratory indicators of adverse reactions to foods. The Collected Papers of the International College of Applied Kinesiology 2006;35–44.
41 Bachmann LM, Puhan MA, ter Riet G, Bossuyt PM: Sample sizes of studies on diagnostic accuracy: literature survey. BMJ 2006;332:1127–1129.
42 Harper R, Reeves B: Reporting of precision of estimates for diagnostic accuracy: a review. BMJ 1999; 318:1322–1323.
43 Hass M, Peterson D, Hoyer D, Ross G: The reliability of muscle testing response to a provocative vertebral challenge. J Manipulative Physiol Ther 1993; 5:95–100.
44 Lawson A, Calderon L: Interexaminer agreement for applied kinesiology manual muscle testing. Percept Mot Skills 1997;84:539–546.
45 Peterson KB: A Preliminary inquiry into manual muscle testing response in phobic and control subjects exposed to threatening stimuli. J Manipulative Physiol Ther 1996;19:310–316.
46 Scopp AL: An experimental evaluation of kinesiology in allergy and deficiency disease diagnosis. J Orthomol Psychiatry 1979;7:137–138.
47 Staehel HJ, Koch MJ, Pioch T: Double-blind study on materials testing with applied kinesiology. J Dent Res 2005;84:1066–1069.
48 Chung A-L, Shin E-J, Kim K: Reliability of the kinesiologic occlusal position. Int J Appl Kinesiol Kinesiol Med 2005;20:6–10.
49 Jacobs GE: Applied kinesiology: an experimental evaluation by double blind methodology. J Manipulative Physiol Ther 1981;4:141–145.
50 Friedman MH, Weisberg J: Applied kinesiology – double-blind pilot study. The J Prosthet Dent 1981; 45:321–323.
51 Monti DA, Sinnott J, Marchese M, Kunkel EJS, Greeson JM: Muscle test comparisons of congruent and incongruent self-referential statements. Percept Mot Skills 1999;88:1019–1028.
52 Triano JJ: Muscle strength testing as a diagnostic screen for supplemental nutrition therapy: a blind study. J Manipulative Physiol Ther 1982;5:179–181.
53 Gregory WM, Mills SP, Hamed HH, Fentiman IS: Applied kinesiology for treatment of women with mastalgia. Breast 2001;10:15–19.
54 Lewith GT, White PJ, Kaptchuk TJ: Developing a research strategy for acupuncture. Clin J Pain 2006; 22:632–638.
55 Paterson C, Dieppe P: Characteristics and incidental (placebo) effects in complex interventions such as acupuncture. BMJ 2005;330:1202–1205.
56 Conable K, Corneal J, Hambrick T, Marquina N, Zhang J: Electromyogram and force patterns in variably timed manual muscle testing of the middle deltoid. J Manipulative Physiol Ther 2006;29: 305–314.