14. Englische Kurzzusammenfassung

Interrater reliability in evaluating the body language based on the Fascial Distortion Model (FDM)

Stefan Anker

Vienna School of Osteopathy, Danube University Krems –
Centre for Traditional Chinese Medicine and Complementary Medicine

Objective: To investigate the inter-rater reliability of the visual FDM-Body language diagnosis and to observe whether the inter-rater reliability is influenced by profession, experience of verbal statements of patients.

Methods: Patient videos were presented using online questionnaires, in which most patients described their physical complaints nonverbally, to twenty-five osteopaths from Germany, Austria and Switzerland. The testers diagnose of fascial distortions is based on the FDM guidelines. The results are statistically analyzed using the kappa index according to Siegel and Castellan.

Results: The results are statistically analyzed using the kappa index, according to Siegel and Castellan, and show acceptable levels of reliability (k = 0.51). Here, the results of osteopaths with either medical or non-medical backgrounds are significantly better than the results of osteopaths who are physical therapist by trade. Verbal statements that are made by patients parallel to the non-verbal communication do have a significant impact on reliability, in contrast to the duration of the experience, of the tester with FDM, which does not.

Conclusion: Seen against the background of a lack of information about validity, this study presents an acceptable inter-rater reliability of visual FDM diagnosis. These results concerning the influence of verbal statements can only be seen as an encouragement for further research.

Introduction

The reliability of diagnostic tests is of great value in osteopathy. Only with a reliable diagnosis the appropriate treatment for a patient can be found or determined whether the patient generally responds to treatment (Lucas and Bogduk 2011). This complies with the requirements of the evidence-based medicine (EBM), which constitute reliable tests as a basic element for comprehensible behaviour (Sackett et al. 1996). Osteopathy predominantly uses manual test methods for diagnosis (Biberschick 2011). Reliability studies in the field of manual test methods show more or less acceptable results, as revealed by review articles (Seffinger et al. 2004 Hollerwöger 2006, Stochkendahl et al. 2006, Haneline and Young 2009). They also show quality problems that undermine the validity of the processed studies. Many reliability studies lack an independent reference standard and neglect therefore statements about the validity of the particular test method (Seffinger et al. 2004, Hollerwöger 2006, Stochkendahl et al. 2006). A further problem is the lack of standardization of the examination process (Stochkendahl et al. 2006, Hollerwöger 2006). In addition measurements are performed on non-representative study populations (Hollerwöger 2006, Stochkendahl et al. 2006).

In contrast to this manual approach of osteopathy the Fascial Distortion Model (FDM) supports a visual approach to clinical diagnosis, which is applied by a large number of osteopaths in practice.

The FDM is an anatomical model offers the cause of physical symptoms in six potential anatomical changes of the shape of the fascia. Through the concept of FDM, developed by osteopathic physician and emergency doctor Stephen Typaldos, a number of diseases in orthopaedics, internal medicine and neurology can be explained. As a result specific treatments can be applied and effects determined.

FDM presents a specific approach to diagnosing fascial distortions, which strongly correlates to subjective verbal and nonverbal communication of the patient. Partly the diagnosis in the FDM is based on a type of visual diagnostics, as suggested by Typaldos (1999, 2002). Fascial distortions are presented through patients indicating them on their own body. Typaldos (2002) describes this process as body language, referring to consistent, unconscious movements or postures. From a scientific point of view body language is bound to be associated with the term gesture. Gesture is defined by the predominant movements of the arms and head (Duden online 2011) and by a distinct temporal process (Kendon 1996, McNeill 1992) with an initial starting, peak (stroke) and final end sequence of the gesture. These

criteria are of no importance in FDM. Thus, scientific information from gesture science can only be used with reservations for FDM.

Reliability studies on visual diagnosis of gestures and body language are incidental in medical databases. Some comparable studies in the field of cardiology (Edmondstone 1995, Albarran et al. 2000, Marcus et al. 2007) show that gesture obviously holds certain information and reference matter. In the cited studies (Edmondstone 1995, Albarran et al. 2000, Marcus et al. 2007) no strong evidence of acceptable reliability could be shown. In contrast to these studies visual diagnosis in FDM lacks independent reference standards. Fascial distortions and their diagnosis through FDM body language are part of the model-like presentation of FDM and are currently of no consequence outside this system in general medicine. Since the FDM-diagnosis was not investigated regarding its inter-rater reliability, no statements concerning the validity of FDM tests are possible.

Generally, FDM distinguishes six fascial distortions. These are distinguished by their formation mechanism, location and quality of complaints, their extend and, above all, through their body language by which they can be encoded in the visual FDM diagnosis (see Table 1).

Fascial distortions	Definition, complaints and verbal description	Body language
Triggerbands	Deformation of banded fascia, which is caused by twisting, cutting, tearing and wrinkling of fibres; patients mostly describe burning or aching pain along a subjectively defined anatomical process; movement restrictions, loss of strength or swelling are other possible indications of a triggerband	Stroking along painful/tender parts with several fingers
Herniated triggerpoints	Protrusion of deeper tissue through overlying fascial layer; is found almost exclusively on the core; causing dull pain, also in the region around the herniation; possible restriction of movement of the adjacent joints	Press with fingers, thumbs or knuckles in the protrusion area
Continuum distortions	Distortion of the transition zone between bone and ligament, tendon or other fascial structures; pain or restricted movement in one axis as a result	Point a finger at one or more painful spots near bones

Fascial distortions	Definition, complaints and verbal description	Body language
Folding distortions	3-D deformation of the joint covering fascia, intramuscular septum or the interosseous membrane; symptoms are painful movement, pain in the lower levels, feeling of instability, causing slight movement restrictions; distinction between Unfolding and Refolding distortions which can be sometimes seen in different body language	Rubbing or pulling of the joint with the hand or hands; hand is put (depending on availability) on the affected area; fists pressed into the back, supporting the back or relieving it with the help of the hands; fingers stroking parallel to the joint line is a reference of a refolding distortion; pressing several fingers between the muscle venter or bones, pulling or tugging on the muscles is presented in folding distortions of interosseous membranes and inter-muscular septa; folding distortions are often displayed dynamically with repetitive movement of the affected body part
Cylinder distortions	Deformation of the spiral turns of the circular fascia; possible symptoms such as paraesthesia, cramps, tremors, paralysis, etc.; there is often a large discrepancy between shutter release and the present complaints; the pain is felt "in depth"; from outside usually not palpable	Massaging or stroking of the soft tissues, wiping hand movements, on occasion the body part is not touched at all
Tectonic fixation	Loss of slippage of fascial surfaces; causes little pain, however, limited mobility on all levels and the feeling of stiffness	Holding and massaging of the joint; patients try to mobilize themselves with own strength; shoulder abduction is only possible with anterior elevation, in the abdominal position the arm cannot be stretched forward horizontally without the upper body lifting off the pad; in case of stiffness in the hip placing palms on both christae iliac; holding and rubbing of the joint

Table 1: Characteristics of the six fascial distortions and their body language (compare Typaldos 1999, 2002)

In FDM diagnosis body language is entwined with verbal description of the complaints of the patients, movement and exercise tests and palpation, which is mainly used in the FDM for differential diagnosis.

The present study examines how reliable osteopaths evaluate body language in terms of the Fascial Distortion Model. Additionally, the influence of the profession of the osteopath (physician, medical practitioner, physiotherapist), experience with FDM and further verbal statements of the patients on the inter-rater reliability should be investigated. Here, the points of criticism of the quality of reliability studies (see above) will be considered in the study design.

Methods

The study was designed as a video-assisted online questionnaire.

Videos

50 short videos of patients describing their complaints non-verbally are filmed, qualitatively retouched and divided into sequences in the practice of the investigator. First, 23 of those videos are randomly selected. Another three of these videos, which appeared twice in the questionnaire with and without sound are randomly chosen. Further two videos that contain no diagnosable body language are added. In result 28 assessable videos are integrated in the online questionnaire.

The online questionnaire allows a high degree of standardization of the examination process. It consists of 33 questions. First, the testers are informed about the investigation process, the conditions for the execution of the questionnaire and criteria of how to answer the questions. It is a single-choice questionnaire, and once the testers have watched the video, they must decide which of the six fascial distortions can be diagnosed on the basis of FDM criteria. Additional answers are possible: "Keine Fasziendistorsion (No fascial distortion)" if no fascial distortion is visible, or "Weiß nicht (Do not know)" if the question cannot be answered. The submission of a response is necessary to continue the questionnaire. Subsequent alteration of the responses is not possible. The first three questions refer to the testers' original profession, their training in osteopathy and duration of the experience with FDM. Two test videos are shown to the testers to practice the examination process. Their results are not included in the statistics. The 28 questions of the study (v6 to v33) are to be answered according to the presented method. Three videos, first without then with sound, are shown to investigate the impact of verbal statements alongside with non-verbal body language have on reliability.

Testers

The testers must fulfil following criteria: trained osteopaths, a FDM training, according to the guidelines of the EFDMA (European Association FDM 2010), and sufficient knowledge of the German language. 79 therapists (doctors, physiotherapists, medical practitioners) from Germany, Austria and Switzerland are invited per email, and 25 of those meet all the necessary criteria. There are 5 female and 20 male testers who will answer the questionnaire in full. This data then is included in the statistics.

Statistics

For the evaluation of reliability the kappa index (k index), according to Siegel and Castellan (1988), is used. In contrast to Cohen's kappa, which is limited to two investigators, this is a generalized form of the kappa index. This means that the compliance of all investigators can be expressed in one index. Another advantage is the possibility in the case of partially absent data to calculate the index. In the case of this study the possible answer „Weiß nicht (Do not know)" is interpreted as lack of response and therefore makes this statistical method of calculation necessary.

The proposed textual interpretation according to Landis and Koch (1977) can be applied for the kappa index of Siegel and Castellan (1988), (see Table 2).

k < 0,20	poor	schwach
0,20 < k < 0,40	adequate	leidlich
0,40 < k < 0,60	moderate	mittelmäßig
0,60 < k < 0,80	substantial	beträchtlich
0,80 < k < 1,00	almost ideal	fast ideal

Table 2: Interpretation of reliability scores by Landis and Koch (1977)

According to Fjellner et al. (1999) values by at least 0.4 are considered to be an acceptable indicator of inter-individual reliability.

In addition to the generalized kappa index, all of the investigator-pairs (according to Siegel and Castellan 1988) are calculated in order to get an overview of the spread of the kappa indices. This data also forms the basis for further comparisons of the various subgroups.

Subsequently, through variance analysis and significance tests (independent samples t-tests or Wilcoxon tests) is examined to test whether there are significant differences between the various sub-groups (see Table 3) in terms of expectation value.

Osteopaths		
A1. Doctors	A2. Physiotherapists	**ANOVA**
A1. Doctors	A3. Medical practitioners	
A2. Physiotherapists	A3. Medical practitioners	
A4. Long experience (>3 years)	A5. Short experience (1-3 years)	t-Test
A6. Long experience (>2 years)	A7. Short experience (1 and 2 years)	t-Test

Table 3: Analysis of variance (ANOVA) and significance tests (Independent samples t-Test. .. t-test, Wilcoxon Wilcoxon test ...) as to test the expected value of the k-index compared among sub-groups.

To examine whether verbal statements by the patient influence the results three videos, with and without sound, are presented to the testers. The reviews of these videos present the base of the data for the calculation of the k-indices for the corresponding pairs of the investigators.

These in turn were compared with each other by means of significance tests (independent variable "sound"), where the kappa indices were used for substitution of the negative value of zero, since their results are widely comparable to the generalized kappa index.

Furthermore, the results of the individual questions were specifically investigated using the qualitative criteria of gesture studies in order to obtain possible correlations between the inter-rater reliability for individual videos and their nature and content.

Results

The results relating to this research can be summarized as follows (see Table 4; page 118):

The overall index is k = 0.51 (according to Siegel and Castellan, 1988) and it is significantly different from zero (Z = 53.42, p <0.0001). The mean average of all paired comparisons from the reviews of Osteopaths calculated k indices adds up to k = 0.57 ± 0.08 (95% CI: 0.54 to 0.60).

Consequently one can speak of an acceptable inter-rater reliability (after Fjellner et al. 1999) of the FDM visual diagnosis in the context of this study.

The mean k-index of the physical therapists in the sample is significantly different both from that of doctors (t = 5.180, df = 68.2, p <0.0001) and that of the medical practitioners (t = 3.040, df = 7.8, p = 0.02). The mean k-value of the physical therapist is lower by 0.14 or 0.13 compared to the one of the control groups. There is no significant difference (difference: k = 0.01, t = 0.1900, df = 8.2, p = 0.85) between the reliability of the opinions of doctors and that of the practitioners. Doctors reach an average index of k-k = 0.57 ± 0.13 (95% CI: 0.53 to 0.61), the Therapy k = 0.56 ± 0.10

Fascial distortion	1 - Triggerband TB	2 - Herniated triggerpoint HTP	3 – Continuum distortion CD	4 – Folding distortion FD	5 – Cylinder distortion CyD	6 – Tectonic fixation TF	7 - No fascial distortion		Valid answers (n)	Do not know - Missings
V6	1	2	***11***	***11***	0	0	0		25	
V7	0	5	***15***	3	0	0	0		23	2
V8	2	5	***13***	2	0	0	0		22	3
V9	***25***	0	0	0	0	0	0		25	
V10	***24***	0	0	0	0	0	0		24	1
V11	4	0	0	4	***17***	0	0		25	
V12	3	0	0	***18***	3	1	0		25	
V13	0	7	0	0	0	***13***	3		23	2
V14	0	0	0	0	***25***	0	0		25	
V15	1	0	***23***	0	1	0	0		25	
V16	5	0	1	***16***	0	0	0		22	3
V17	1	0	2	***21***	0	0	0		24	1
V18	***22***	0	1	0	1	0	0		24	1
V19	5	***12***	2	3	1	1	0		24	1
V20	***25***	0	0	0	0	0	0		25	
V21	***16***	0	0	0	9	0	0		25	
V22	2	0	1	***21***	1	0	0		25	
V23	***17***	2	1	3	2	0	0		25	
V24	***23***	0	0	0	2	0	0		25	
V25	0	***24***	0	0	0	0	0		24	1
V26	0	0	***25***	0	0	0	0		25	
V27	0	0	0	***13***	10	2	0		25	
V28	***22***	0	0	0	3	0	0		25	
V29	1	0	0	4	***19***	0	0		24	1
V30	6	7	0	***8***	3	1	0		25	
V31	0	0	0	***10***	5	0	1		16	9
V32	1	0	0	0	***18***	0	3		22	3
V33	8	**9**	0	6	1	1	0		25	

Table 4: Frequency of diagnoses by the 25 osteopaths for each individual patient. (bold & italic ... most likely diagnosis, dark green mark:> 90% (n> 22) line, light green ...> 80% (n> 20), light yellow ...> 70% (n> 17) and orange. ..> 60% (n> 15).

(95% confidence interval: 0 0.48-0, 64) and the physiotherapists k = 0.43 ± 0.11 (95% CI: 0.40 to 0.46).

Even without considering the examiners 2 and 7, who have significantly poorer agreement with the other physical therapists, the reliability lies, however, at k = 0.49 ± 0.08 (95% CI: 0.46 to 0.52), which is significantly lower than that of the doctors (Wilcoxon W = 687.5, p = 0.01).

On average, there was no significant difference in reliability of those osteopaths with more than two years experience with the FDM and those with less than three years experience (difference: k = 0.03, t = 1.097, df = 17.5, p = 0, 28). Even taking into account a different threshold for high and low experience no significant differences were observed. The difference between the mid-level reliability of those osteopaths with more than three years experience with the FDM and those with less than four years experience is k = 0.01 (t = 0.091, df = 53.0, p = 0.93).

The comparison of the average of the k-indices, which were calculated from the assessments of three videos with sound and those without sound, shows a significant difference (Wilcoxon W = 40598.5, p = 0.03). With sound, the average value amounts to k = 0.24 ± 0.30 (95% CI: 0.20 to 0.27), without sound k = 0.17 ± 0.26 (95% CI: 0.14 - 0.20).

Discussion

The visual diagnostics seems to be a reliable test for the diagnosis of fascial distortions according to the FDM. The acceptable results of this investigation throw a positive light on this diagnostic approach, applied by osteopaths in practice.

This is the first study that examines the inter-rater reliability of the FDM-body language and there is no independent reference standard for the diagnosis of fascial distortions. These factors must be mentioned in the evaluation of the results.

Based on the qualitative analysis of individual results of each video a greater spreading of results of the videos with lower tester agreement was observed. Interestingly, with the decline of the agreements results accumulated which cannot be related to the criteria for the diagnosis of specific distortion types.

In the same videos also a build up of incomplete temporal gesture sequence in the sense of Kendon (1996) and McNeill (1992) is evident. That could make it difficult

for the tester to detect the peak gesture followed by a smaller agreement level. To what extend these observations relate to the structure of the questionnaire in its present form, as well as the video section or the lack of FDM criteria for body language in itself, remains unclear at this point.

Further the results show that from 700 submitted answers only in 35 cases the testers responded with the options „Keine Fasziendistorsion (No fascial distortion)“ or „Weiß nicht (Do not know)", although at least two of the videos showed no evidence of encodable body language in terms of FDM. This raises the question of how the criteria of body language vocabulary of FDM are defined.

The different results of inter-rater reliability with respect to the testers' profession challenge the findings of Seffinger et al. (2004) and Stochkendahl et al. (2006), which cannot confirm any connection between the testers and the inter-rater reliability in their reviews.

A possible explanation for this statistical difference might be a different composition of the individual test groups in relation to their experience or gender; these criteria are not included in the statistics. Due to the low variance in the range of volume of experience of the testers, the impacts of these results in this study are unlikely. Through a more consistent distribution of experience the results for the investigation of the relationship between experience and inter-rater reliability could be affected.

In this study no correlation between the experience of the tester and the inter-rater reliability can be found. This may be related to the selection of testers according to their experience. A more consistent distribution of experience with FDM among the testers could alter the results. This result correlates with the reviews of Seffinger et al. (2004), Stochkendahl et al. (2006) and Haneline and Young (2009), which cannot establish a general relationship of reliability and experience.

There is a significant impact of verbal statements on the inter-rater reliability. Though the limiting to three video pairs, these results are to be examined and thus must be seen as a stimulus for further research. Here, a much larger sample group would be necessary to give more weight to these results. In the case of this study the consistent diagnostic results are obtainable through verbalization.

Looking at the results of this study following can be criticised. A higher number of videos in the questionnaire, with significant differences of filmed patients regarding

gender, symptoms, affected body regions and so on, could better represent the wide variety of FDM body language and thus reinforce the significance of the results and their application in practice. The separating of the production, especially the post processing, of the videos from the investigator should be considered to obtain a more blinding effect in this context.

The videos in this study are partially edited to reduce possible internal errors. The videos are also faded in and out, which could create an artificial gestural climax of the individual videos that may not be comparable to the "natural" peak of the gesture. This should be considered in future studies solving the problem through editing or longer video clips. The potential increase of the internal errors in single choice study design should be considered in this context.

The video setting within an online questionnaire shows no possibility of criticism or ability to enquire upon lack of clarity of the seen body language, which doesn't correspond with the practical approach.

Because of technical factors, it is not possible to prevent a repeated viewing of the videos by the testers. Although the reproducibility itself is an advantage, when used for video recording in reliability analysis (Haidet et al. 2009) and in measurement accuracy (Sommerfeld 2006), it should be noted that this is not relevant in this code of practice where body language in the same manner is mostly shown only once. Further critique of the structure of the online questionnaire it is whether the testers can actually distinguish between the possible answer „Keine Fasziendistorsion (No fascial distortion)" and „Weiß nicht (Do not know)". A wider awareness of the theoretical background of these choices and associated research questions would be necessary to generate corresponding sensitivity on part of testers.

The main problem in this study, as with most reliability studies in the field of osteopathic manual medicine, is the discrepancy between internal and external validity. The reductionist approach in such scientific studies only partially reflects the clinical practice. In the study at hand, the problem is that body language occurs only rarely in isolation, and through the observation within the osteopath-patient communication, it occurs in the context of verbal statements and the patient's medical history. Thus for this study one cannot say anything about the reliability of the FDM-diagnosis in general, but only on their visually evident part that has to be (as described) cross-linked in a particular case with other information from the investigation in order to attain FDM diagnostic approach.

Conclusion

Using a video-assisted online survey, the body language shows acceptable inter-rater reliability in this master thesis. Since this visual diagnostic approach cannot currently be reviewed by an independent reference standard the validity of body language-diagnosis is not clear. These results should be transferred with caution into practice. In particular the use of the test as a singular diagnostic criterion is critical. In the context of the entire FDM-specific findings for the diagnosis of fascial distortions, the results of this study, however, are significant.

Acknowledgement

Thanks to Mag[a]. Claudia Gamsjäger for the care of this master thesis.

Literature

Albarran JW, Durham B, Chappel G, Dwight J, Gowers J (2000): Are manual gestures, verbal descriptors and pain radiation as reported by patients reliable indicators of myocardial infarcation? Preliminary findings and implications. Intensive and Critical Care Nursing, 16(2), 98 – 110

Biberschick M. (2011): „Legt Euch ein Schema zurecht!“: die Routineuntersuchung in der Osteopathie. Master Thesis zur Erlangung des akademischen Grades „Master of Science“ in Osteopathie. Krems. Donau Universität Krems

Duden online. (2011): *http://www.duden.de/*

European FDM Association EFDMA. (2010): FDM-Ausbildung. *http://www.fdm-europe.com/fdm-ausbildung/ausbildung-fdm--curriculum/index.html* (Accessed: Oktober 10, 2010)

Edmondstone WM (1995): Cardiac chest pain: Does body language help the diagnosis? British Medical Journal, 311(7021), 1660 – 1661

Fjellner A, Bexander C, Faleji R, Strender LE. (1999): Interexaminer reliability in physical examination of the cervical spine. J. Manipulative Physiol Ther. 22, 511 – 516

Haidet KK, Tate J, Divirgilio-Thomas D, Kolanowski A, Happ MB. (2009): Methods to Improve Reliability of Video Recorded Behavioral Data. Res Nurs Health, 32 (4): 465 - 474

Haneline MT, Young M. (2009): A review of intraexaminer and interexaminer reliability of static spinal palpation: a literature synthesis. J Manipulative Physiol Ther. 32, 379 – 386

Hollerwöger D. (2006): Methodological quality and outcomes of studies addressing manual cervical spine examinations: a review. Man Ther. ,11, 93-98

Kendon A. (1996): An Agenda for Gesture Studies. Semiotic Review of Books. 7 (3), 8 – 12

Landis, JR, Koch GG. (1977): The measurement of observer agreement for categorical data. Biometrics. 33, 159 – 174

Lucas N, Bogduk N (2011): Diagnostic reliability in osteopathic medicine. International Journal of Osteopathic Medicine,1, 1 - 5

Marcus GM, Cohen J, Varosy PD, Vessey J, Rose E, Massie BM, Chatterjee K, Waters D. (2007): The utility of Gestures in Patients with Chest Discomfort. The American Journal of Medicine, 120, 83 – 89

McNeill D. (1992): Hand and mind: What gestures reveal about thought. Chicago: The University of Chicago Press

Sackett DL, Rosenberg WMC, Gray Muir JA, Haynes RB, Scott Richardson W. (1996): Evidence based medicine: what it is and what it isn't. British Medical Journal, 312, 71

Seffinger MA, Najm WI, Mishra SI, Adams A, Dickerson VM, Murphy LS, Reinsch S (2004): Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. Spine, 1, 29, 413 – 25

Siegel S, Castellan NJ Jr. (1988): Nonparametric statistics for the behavioural sciences. Boston: McGraw-Hill

Sommerfeld P (2006): Methodologie. Kleine und hoffentlich kurzweilige Einführung in den Sinn und Unsinn wissenschaftlichen Arbeitens auf der Grundlage quantitativer Methoden für OsteopathInnen. 3. Überarbeitete Version.

Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, Adams A. Bronfort G (2006): Manual Examination of the Spine: A Systematic Critical Literature Review of Reproducibility. Journal of Manipulative and Physiological Therapeutics, 29, 6, 475 – 85

Typaldos S. (1999): Orthopathische Medizin. Die Verbindung von Orthopädie und Osteopathie durch das Fasziendistorsionsmodell. Kötzing/Bayer. Wald: Verlag für Ganzheitliche Medizin Wühr

Typaldos S. (2002): FDM – Clinical and Theoretical Application of the Fascial Distortion Model within the Practice of Medicine and Surgery. Bangor: Eigenverlag