

University of Groningen

Interexaminer Agreement and Reliability of an Internationally Endorsed Screening Framework for Cervical Vascular Risks Following Manual Therapy and Exercise

de Best, Rogier F; Coppieters, Michel W; Trijffel, Emiel; Compter, Annette; Uyttenboogaart, Maarten; Bot, Joost C; Castien, Rene; Pool, Jan J M; Cagnie, Barbara; Scholten-Peeters, Gwendolyne G M

Published in:
Physical Therapy

DOI:
[10.1093/ptj/pzab166](https://doi.org/10.1093/ptj/pzab166)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

de Best, R. F., Coppieters, M. W., Trijffel, E., Compter, A., Uyttenboogaart, M., Bot, J. C., Castien, R., Pool, J. J. M., Cagnie, B., & Scholten-Peeters, G. G. M. (2021). Interexaminer Agreement and Reliability of an Internationally Endorsed Screening Framework for Cervical Vascular Risks Following Manual Therapy and Exercise: The Go4Safe Project. *Physical Therapy*, 101(10), [166]. <https://doi.org/10.1093/ptj/pzab166>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Interexaminer Agreement and Reliability of an Internationally Endorsed Screening Framework for Cervical Vascular Risks Following Manual Therapy and Exercise: The Go4Safe Project

Rogier F. de Best, MSc¹, Michel W. Coppieters, PhD^{1,2}, Emie van Trijffel, PhD^{3,4}, Annette Compter, MD⁵, Maarten Uyttenboogaart, MD⁶, Joost. C. Bot, MD⁷, Rene Castien, PhD¹, Jan J.M. Pool, PhD⁸, Barbara Cagnie, PhD⁹, Gwendolyn G.M. Scholten-Peeters, PhD^{1,*}

¹Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, The Netherlands

²Menzies Health Institute Queensland, Griffith University, Brisbane and Gold Coast, Australia

³SOMT University of Physiotherapy, Amersfoort, The Netherlands

⁴Department of Experimental Anatomy, Vrije Universiteit Brussel, Brussels, Belgium

⁵Department of Neuro-Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands

⁶Department of Neurology and Radiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

⁷Department of Radiology, Amsterdam University Medical Center, The Netherlands

⁸Research Group Lifestyle and Health, HU University of Applied Sciences Utrecht, Utrecht, The Netherlands

⁹Department of Rehabilitation Sciences, Ghent University, Ghent, Belgium

*Address all correspondence to Dr Scholten-Peeters at: g.g.m.scholten-peeters@vu.nl

Abstract

Objective. Clinicians are recommended to use the clinical reasoning framework developed by the International Federation of Orthopaedic Manipulative Physical Therapists (IFOMPT) to provide guidance regarding assessment of the cervical spine and potential for cervical artery dysfunction prior to manual therapy and exercise. However, the interexaminer agreement and reliability of this framework is unknown. This study aimed to estimate the interexaminer agreement and reliability of the IFOMPT framework among physical therapists in primary care.

Methods. Ninety-six patients who consulted a physical therapist for neck pain or headache were included in the study. Each patient was tested independently by 2 physical therapists, from a group of 17 physical therapists (10 pairs) across The Netherlands. Patients and examiners were blinded to the test results. The overall interexaminer agreement, specific agreement per risk category (high-, intermediate-, and low-risk), and interexaminer reliability (weighted κ) were calculated.

Results. Overall agreement was 71% (specific agreement in high-risk category = 63%; specific agreement in intermediate-risk category = 38%; specific agreement in low-risk category = 84%). Overall reliability was moderate (weighted κ = 0.39; 95% CI = 0.21–0.57) and varied considerably between pairs of physical therapists (κ = 0.14–1.00).

Conclusion. The IFOMPT framework showed an insufficient interexaminer agreement and fair interexaminer reliability among physical therapists when screening the increased risks for vascular complications following manual therapy and exercise prior to treatment.

Impact. The IFOMPT framework contributes to the safety of manual therapy and exercise. It is widely adopted in clinical practice and educational programs, but the measurement properties are unknown. This project describes the agreement and reliability of the IFOMPT framework.

Keywords: Adverse Events, Carotid Artery, Physiotherapy, Reproducibility, Spinal Manipulation, Vertebral Artery

Received: July 5, 2020. Revised: March 25, 2021. Accepted: May 27, 2021

© The Author(s) 2021. Published by Oxford University Press on behalf of the American Physical Therapy Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

Neck pain and headache are 2 conditions commonly treated by physical therapists.^{1–3} Systematic reviews have shown small to moderate levels of evidence for the effectiveness of cervical manual therapy and exercise for patients with neck pain^{4–6} or headache.^{7,8} However, because of associated serious adverse events, such as cerebral ischemia and even death, there is discussion about the safety of manual therapy for the cervical spine.^{9–11} Potential hypotheses for serious adverse events are alterations of hemodynamics in the cervical arteries, arterial dissection, or atherosclerosis,^{9,12,13} and missed preexisting vascular pathologies that mimic musculoskeletal neck pain or headache,^{14,15} including arterial dissection and atherosclerosis. The prevalence of vertebral artery dissection in the general population is 0.75–2.90 per 100,000 people, and the prevalence of vascular dissections associated with manual therapy is estimated between 0.40 per 100,000 to 5.00 per 100,000. There are no data on the prevalence of complications following exercise.^{16,17}

The International Federation of Orthopaedic Manipulative Physical Therapists (IFOMPT) developed a framework based on the best available evidence to provide guidance regarding assessment of the cervical spine and potential of cervical artery dysfunction prior to manual therapy, with special attention to interventions at end-range positions of the cervical spine.¹⁸ The multivariable character of this clinical reasoning framework is in accordance with physical therapist practice in which diagnostic conclusions are based on multiple rather than single tests.¹⁹ The IFOMPT framework is intended to improve early identification of patients with neck pain or headache with an increased risk of developing serious adverse events following cervical manual therapy or exercise.

Although the framework is widely adopted in clinical practice and educational programs,²⁰ the reproducibility and accuracy of the IFOMPT framework have not yet been assessed. Agreement, specific agreement, and reliability provide an insight into reproducibility. Agreement concerns how well outcomes from different examiners agree. It is expressed in terms of observed agreement. The proportion of specific agreement distinguishes agreement on positive or negative scores.²¹ Reliability concerns how well patients can be distinguished from each other despite measurement errors and is expressed as a coefficient varying from 0 to 1.²²

At present, there is lack of information about the measurement properties of the IFOMPT framework. Consequently, this study aimed to estimate the interexaminer agreement and reliability of the risk estimation of vascular complications following manual therapy and exercise prior to treatment among physical therapists.

Methods

Design

We conducted an interexaminer agreement and reliability study, reported here according to Guidelines for Reporting Reliability and Agreement Studies.²³ The Medical Ethics Review Committee of the Amsterdam University Medical Centre (Location VUmc) approved the study protocol (METC-2017.086).

Participants

Consecutive patients with neck pain and/or headache were recruited from primary care physical therapist practices in The

Netherlands between July 2017 and March 2019. Patients were eligible if they were at least 18 years old, consulted a physical therapist for neck pain or headache, and had sufficient knowledge of the Dutch language. All patients signed informed consent prior to inclusion in the Go4Safe project.

Examiners

Thirty-four physical therapists were recruited via social media, the Internet, newsletters of the Royal Dutch Society of Physiotherapy, and physical therapy networks. Thirty-two of these physical therapists had a master's degree in manual therapy with at least 2 years of relevant clinical experience and were eligible to participate. Characteristics of the physical therapists are summarized in [Table 1](#). All physical therapists attended a 3-hour refresher course on how to conduct the clinical reasoning framework and how to interpret the risk of vascular complications.

During the refresher course, the early and late clinical manifestations of cervical artery dysfunction, such as the description of pain as “unlike any other,” and the risk factors according to the IFOMPT framework were highlighted.

Interactive case studies and clinical reasoning tasks focused on the estimation of the probability of serious vascular pathology or vascular complications following manual therapy and exercise. Participants practiced the clinical tests for the cranial nerves, palpation of the carotid artery, and blood pressure measurements on each other. Blood pressure was measured according to the guidelines from the National Institute for Health and Care Excellence.²⁴ The training was delivered by 2 experienced instructors of the Masters of Science program in Manual Therapy at SOMT University of Physiotherapy. The instructors had 6 and 14 years of teaching experience in manual therapy, and 10 and 37 years of relevant clinical experience.

To ensure blinding of the patients and examiners, no information was given about the outcome of the test. Seventeen physical therapists (53%) actually included patients in the study. These formed 10 separate pairs from 10 different physical therapist practices. There were no significant differences in age, sex, or experience in physical therapy or manual therapy between physical therapists who provided patients compared with those who did not provide patients for the study ([Tab. 1](#)).

Test Procedures

Eligible patients were examined by their treating physical therapist using the IFOMPT framework.¹⁸ Subsequently, a patient interview and, if necessary, clinical testing of the cervical blood vessels (ie, evaluation of blood pressure and palpation of the carotid artery) and/or cranial nerves was performed. Afterwards, each physical therapist estimated the risk of a vascular complication following manipulation, mobilization, and exercise therapy as “low,” “intermediate,” or “high.”¹⁸ The [Figure](#) summarizes the flow of the IFOMPT framework.

When the physical therapist expected a vascular origin (eg, dissection, atherosclerosis) for the symptoms, or when multiple risk factors for vascular complications (eg, stroke) following manual therapy were present, a high-risk score was given. The risk estimation (“high risk,” “intermediate risk,” or “low risk”) was based on a thorough clinical reasoning process in which risk factors for vascular complications, the results of patient interview, and physical examination were weighted by the physical therapist. No

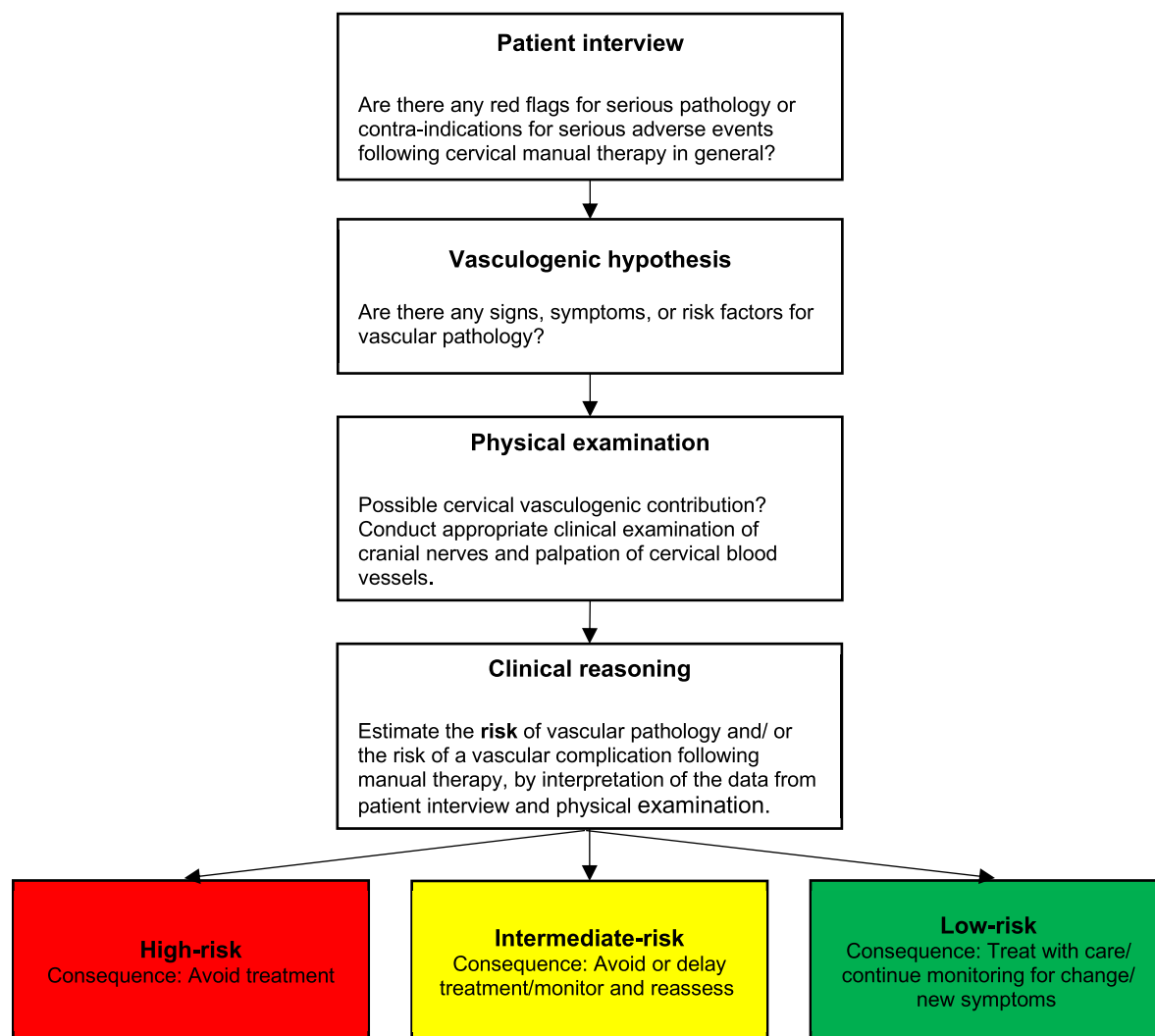


Figure. Flow of the IFOMPT (International Federation of Orthopaedic Manipulative Physical Therapists) framework.

Table 1. Characteristics of the Physical Therapists (N = 34)^a

Characteristics	Active PT (n = 17)	Nonactive PT (n = 17)	P
Sex (F), n (%)	5 (29%)	3 (18%)	.69 ^b
Age, y, median (IQR)	36 (31–47)	35 (31–44)	.55 ^c
Experience in physical therapy	12 (8–23)	12 (8–19)	.55 ^c
Experience in manual therapy	7 (4–16)	5 (2–12)	.21 ^c

^aActive PT = physical therapists who included patients in the study; nonactive PT = physical therapists who did not include patients in the study. F = female; IQR = interquartile range. ^bSignificant level of difference between groups using Fisher exact. ^cSignificant level of difference between groups using Wilcoxon rank sum test.

cutoff points for “low,” “intermediate,” or “high” risk were set. The benefits were not considered. Within 1 week, a second physical therapist performed the examination in the same physical therapist practice. Patients and examiners of the second test did not receive information about the outcomes of the first test. In addition to the patient interview and clinical examination by the physical therapists involved in the study, patients received an online questionnaire from the research team to collect demographic data and clinical characteristics (eg, potential risk factors for vascular pathology, neck pain intensity, and additional symptoms). The

physical therapists recorded all adverse events during or after the examination.

Sample Size

A sample size calculation was performed in R-Studio (Version 1.1.453; R-Studio, Boston, MA, USA). Using the “KappaSize” package, 79 participants were needed for this study. This was based on 2 examiners, 3 risk categories, a lower limit of the 95% CI of weighted $\kappa = 0.60$, and an anticipated prevalence of 0.10 in the high-risk category, 0.20 in the intermediate category, and 0.70 in the low-risk category. The estimated

prevalences were based on a pilot study in primary care physical therapist practices.

Statistical Analysis

“Patients” and “physical therapists” characteristics were summarized using descriptive statistics. Data were checked for normality using visual inspection of a histogram, a box-plot, and tested with the Shapiro-Wilk test. Means and SDs were described if numerical data were normally distributed, and median and interquartile range when not normally distributed. In case of categorical data, absolute and relative frequencies were presented.

Three-by-three tables were constructed in which the results of the paired examiners were plotted against each other. To express agreement, we calculated percentage overall agreement and specific agreement. Specific agreements were calculated for each risk category separately.²² To calculate reliability, we used linear weighted κ .

A combination of these coefficients provides a more detailed impression of the degree of agreement and reliability.²³ Agreement and reliability values were calculated for the total group and per examiner pair (2 physical therapists from the same or adjacent physical therapist practice). A specific agreement of at least 75% was considered acceptable.²⁵ For the interpretation of the linear weighted κ , the following criteria were used: almost perfect (0.81–1.00), substantial (0.61–0.80), moderate (0.41–0.60), fair (0.21–0.40), slight (0.00–0.20), and poor (<0.00).²⁶ All data were analyzed with R software version 3.5.1 (CRAN.R-project).

Sensitivity Analysis

According to the IFOMPT framework, the consequences of a high-risk and intermediate-risk score assume no indication for manual therapy at that time. Specifically, the high-risk score implies “Avoid treatment,” and the intermediate risk implies “Avoid or delay treatment/monitor and reassess.” By merging these 2 categories together, the interobserver agreement and reliability of no indication versus indication for manual therapy is assessed, which is relevant for clinical practice.

Therefore, we performed a sensitivity analysis with 2 categories, “Indication for manual therapy” and “No indication for manual therapy,” in which the “high-risk” and “intermediate-risk” groups were combined. Two-by-two tables were constructed in which agreement, specific agreement, and reliability (Cohen κ) were calculated.

Role of the Funding Source

This study was funded by a research grant from the Dutch Association for Manual Therapy (Nederlandse Vereniging voor Manuele Therapie) and was supported by the MSG Science Network Physiotherapy (<https://msg-sciencenetwerk.nl/>). The funders played no role in the design, conduct, or reporting of this study.

Results

A total of 96 patients were included. The median number of patients included per examiner pair was 6 (interquartile range: 4–8). Median age of patients was 53 years (interquartile range: 40–62), and 62 (65%) were female. Sixty-two patients (72%) had neck pain or headache for more than 12 weeks.

Characteristics of the patients are summarized in Table 2. Seventy-five percent of the patients received the retest within 7 days. Median time frame between test and retest was 0 days (interquartile range: 0–8), reflecting that more than half of the patients were tested twice on the same day.

The overall distribution of the results of the first examiner and the second examiner is presented in Supplementary Appendix A. The prevalence of a high-risk category classification for the first examiner was 17% (16/96) and 26% (25/96) for the second examiner.

The overall agreement of the IFOMPT framework was 71% and ranged from 46% to 100% for the different examiner pairs. Overall, specific agreement in the high-risk category was 63%, ranging from 0% to 100%; specific agreement in the intermediate-risk category was 38%, ranging from 17% to 100%; and specific agreement in the low-risk category was 84%, ranging from 57% to 100%. Overall reliability between the examiners was fair (weighted $\kappa = 0.39$; 95% CI = 0.21–0.57). The interexaminer agreement and reliability values are summarized in Table 3.

To address the influence of patients with a time frame of more than 7 days between test and retest on the agreement and reliability results, we performed a post hoc sensitivity analysis without this group. The results changed to: overall agreement 68%, specific agreement in the high-risk category 52%, specific agreement in the intermediate-risk category 39%, and specific agreement in the low-risk category 84%. Overall reliability between the examiners was still fair (weighted $\kappa = 0.31$; 95% CI = 0.10–0.53).

Sensitivity Analysis

By combining the intermediate-risk group with the high-risk group, the prevalence of a high-risk classification increased from 0.17 to 0.36 according to Examiner 1, and from 0.26 to 0.45 according to Examiner 2. The distribution of the sensitivity analysis is presented in Supplementary Appendix B. The overall agreement, specific agreement, and κ values also increased. The κ value increased from 0.39 (95% CI = 0.21–0.57) to 0.60 (95% CI = 0.45–0.77). The interexaminer agreement and reliability values of the sensitivity analysis are summarized in Table 4.

Post hoc Sensitivity Analysis

The main results showed a conspicuously large number of participants in Pair A (n = 50). To estimate the influence of Pair A, we performed a post hoc sensitivity analysis without Pair A. The distribution of the results is presented in Supplementary Appendix C. The agreement and reliability values are presented Table 5. The results revealed agreement and reliability values comparable to the main results, with exception of the specific agreement in the high-risk category. The specific agreement in this category decreased from 63% to 44%.

Adverse Events

No adverse events occurred during the first or second test.

Discussion

This study revealed insufficient agreement and fair reliability of the IFOMPT framework among physical therapists in patients with neck pain or headache who attended primary

Table 2. Self-Reported Characteristics of the Participants (N = 96)^a

Characteristics	Participants
Sex (F), n (%)	62 (65%)
Age, y, median (IQR)	53 (40–62)
BMI, kg/m ² , mean (SD)	26.4 (4.9)
Distribution of symptoms, n (%)	
Occipital headache	1 (1)
Headache around ear and jaw	2 (2)
Mid/upper cervical spine	13 (14)
Lower cervical spine	9 (9)
Mid/upper cervical spine and occipital headache	16 (17)
Mid/upper cervical spine and headache around ear and jaw	1 (1)
Other (eg, lower cervical spine and arm pain)	54 (56)
Pain intensity, NRS, median (IQR)	
Headache	5 (3–7)
Neck pain	5 (3–6)
Cause, n (%)	
Trauma within 30 d ^b	15 (16)
Duration of symptoms, ^c n (%)	
< 6 wk	12 (14)
6–12 wk	12 (14)
> 12 wk	62 (72)
Additional symptoms, ^d n (%)	
Dizziness ^e	34 (38)
Loss of sensibility ^f	31 (39)
Muscle weakness ^g	19 (24)
Possible cranial nerve dysfunction	
Dysphagia	12 (15)
Aphasia	7 (9)
Diplopia	11 (14)
Drop attacks	17 (22)
Other risk factors present, n (%)	
High blood pressure ^b	26 (28)
High cholesterol ^b	24 (27)
History of cardiac arrest ^b	6 (7)
History of CVA ^e	6 (7)
Smoking ⁱ	16 (18)
Alcohol use (>1 unit/d) ^c	7 (8)
Recent infection ^c	8 (9)
Migraine without aura ⁱ	9 (10)
Recent cervical manipulation ^e	50 (56)

^aBMI = body mass index; CVA = cerebrovascular accident; F = female; IQR = interquartile range; NRS = numeric rating scale in the past week; recent infection = recent infection in the past 30 days. ^b4 missing values. ^c8 missing values. ^d17 missing values. ^e7 missing values. ^fLoss of sensibility in arms or legs. ^gMuscle weakness in arms or legs. ^h6 missing values. ⁱ5 missing values. ^j9 missing values.

Table 3. Agreement and Reliability of IFOMPT Framework^a

Examiners	OA	SAH	SAI	SAL	Weighted κ (95% CI)
Overall (N = 96)	71%	63%	38%	84%	0.39 (0.21–0.57)
Pair A (n = 50)	78%	78%	17%	89%	0.47 (0.23–0.72)
Pair B (n = 11)	46%	ND	29%	80%	0.08 (0.00–0.30)
Pair C (n = 8)	50%	50%	40%	57%	0.26 (0.34–0.86)
Pair D (n = 6)	67%	67%	ND	75%	0.40 (0.31–1.00)
Pair E (n = 6)	67%	ND	50%	75%	0.25 (0.00–1.00)
Pair F (n = 5)	100%	0%	100%	100%	1.00 (1.00–1.00)
Pair G (n = 4)	75%	100%	ND	80%	0.60 (0.31–1.00)
Pair H (n = 4)	50%	50%	ND	67%	0.14 (0.00–0.73)
Pair I (n = 1)	100%	ND	100%	ND	ND
Pair J (n = 1)	100%	ND	ND	100%	ND

^aIFOMPT = International Federation of Orthopaedic Manipulative Physical Therapists; ND = no data; OA = overall agreement; SAH = specific agreement of high-risk category; SAI = specific agreement of intermediate-risk category; SAL = specific agreement of low-risk category.

Table 4. Agreement and Reliability of IFOMPT Framework (2 Categories)^a

Examiners	OA	SPA	SNA	κ (95% CI)
Overall (N = 96)	81%	77%	84%	0.60 (0.45–0.77)
Pair A (n = 50)	86%	80%	89%	0.70 (0.49–0.90)
Pair B (n = 11)	82%	83%	80%	0.63 (0.17–1.00)
Pair C (n = 8)	63%	67%	40%	0.25 (0.42–0.92)
Pair D (n = 6)	67%	50%	75%	0.25 (0.6–1.00)
Pair E (n = 6)	67%	50%	75%	0.25 (0.6–1.00)
Pair F (n = 5)	100%	100%	100%	1.00 (ND)
Pair G (n = 4)	75%	80%	67%	0.50 (0.35–1.00)
Pair H (n = 4)	75%	67%	80%	0.50 (0.35–1.00)
Pair I (n = 1)	100%	100%	ND	ND
Pair J (n = 1)	100%	ND	100%	ND

^aIFOMPT = International Federation of Orthopaedic Manipulative Physical Therapists; ND = no data; OA = overall agreement; SNA = negative specific agreement; SPA = positive specific agreement.

Table 5. Agreement and Reliability of IFOMPT Framework (Without Pair A)^a

	OA	SAH	SAI	SAL	Weighted κ (95% CI)
Overall (N = 46)	63%	44%	48%	78%	0.33 (0.10–0.56)

^aIFOMPT = International Federation of Orthopaedic Manipulative Physical Therapists; OA = overall agreement; SAH = specific agreement of high-risk category; SAI = specific agreement of intermediate category; SAL = specific agreement of low risk category.

care. Weighted κ and its 95% CI imply that the IFOMPT framework has a fair to moderate reliability among physical therapists. However, it is important to note that 1 examination pair included 50 patients, which might have influenced the results. Post hoc sensitivity analyses revealed only a small decrease in agreement and reliability, but still insufficient agreement and fair reliability (with boundaries of the CI ranging from slight to moderate).

The sensitivity analyses with 2 categories, “No indication for MT [manual therapy]” and “Indication for MT,” showed an increase of agreement and reliability. Values of κ increased from fair (0.39; 95% CI = 0.21–0.57) to moderate (0.60; 95% CI = 0.45–0.77), and specific agreement improved to an acceptable level (63%–77%). This implies that dichotomization of the outcome categories improves the agreement and reliability of the IFOMPT framework substantially.

To the best of our knowledge, no other studies have assessed the interexaminer agreement or reliability of the IFOMPT framework or other cervical vascular tests. The only information available concerns agreement and is about separate tests in the IFOMPT framework, such as the Sharp-Purser test and premanipulative hold. The Sharp-Purser test is a test for cervical instability, which demonstrates an intraexaminer agreement ranging from 55% to 83%. The premanipulative hold is a sustained hold of the manipulation position for at least 10 seconds and released for 10 seconds.²⁷ The reliability of this test is considered “fair” to “good.”²⁸ However, this needs to be interpreted with caution because of the large variability and wide 95% CIs. None of these studies reported specific agreement values. Other studies of tests advised in the framework merely reported about validity.^{29–31} The physical therapists in this study established the risk of vascular complications following manual therapy and exercise based on the information derived from the patient interview and physical assessment according to the IFOMPT framework. Patients were considered as “high risk” when signs and symptoms suggestive of vascular pathology (eg, dissection or atherosclerosis) were present, or when risk factors for

vascular complications were present. The estimated risk of complications following manual therapy and exercise in this study was much higher (~22%) than the reported prevalence of vascular pathology (ie, vascular pathology related to manual therapy) in the literature.^{16,17,32–34} Moreover, not all “high-risk” patients actually have a vascular pathology or will experience a vascular complication after manual therapy or exercise.

The sample size was calculated with prevalence data for “high,” “intermediate,” and “low” risks derived from a pilot study in primary care physical therapist practices. Post hoc, we recalculated the sample size with the actual parameters and needed a minimum number of 55 patients. Therefore, our sample size (N = 96) was larger than required.

Although we aimed to include consecutive patients, the examiners did not register all eligible patients and reasons why patients refused to participate. Because of the high prevalence in the high-risk category, we assumed that patients who were at risk of vascular complications following manual therapy or exercise were more willing to participate because they received an extra screening by a second independent physical therapist. This implies that our study could contain selection bias, affecting the generalizability of our findings to clinical care. This bias could have overestimated the overall agreement and specific positive agreement values and might have resulted in an increased reliability. The relatively high number of examiner pairs in our study could have negatively affected the agreement values, while also positively influencing the generalizability of our findings in clinical care.

The time interval between the first examination and second examination was set at a maximum of 1 week. However, this could not be met for 23%, and the possibility that these patients received treatment between the 2 examinations cannot be ruled out, nor that this influenced their risk profile. To address the influence of patients with a time frame of more than 7 days between test and retest on the agreement and reliability results, we performed a post hoc sensitivity analysis without this group. This analysis showed results comparable

to the overall results. Therefore, we conclude that patients with a time frame of more than 7 days between tests did not bias our results and that our participants could be considered stable.

The IFOMPT framework proposes a risk-benefit analysis.¹⁸ In a risk-benefit analysis, the risk component is the first and dominant component for safety purposes. The physical therapists in our study were therefore asked only to consider the risk of serious adverse events developing following cervical manual therapy or exercise.

In conclusion, this study found that among physical therapists the IFOMPT framework has an insufficient agreement and only fair reliability in patients with neck pain and/or headache attending primary care when scored into the original 3 categories. However, agreement becomes acceptable, and reliability becomes moderate when “indication for manual therapy” versus “no indication for manual therapy” are considered. Further evidence about the diagnostic accuracy of the risk estimation is necessary to confirm usage in clinical practice. Future research should therefore focus on the diagnostic accuracy of the IFOMPT framework.

Author Contributions

Concept/idea/research design: R.F. de Best, M.W. Coppieters, E. Van Trijffel, A. Compter, J.C. Bot, R. Castien, J.J.M. Pool, B. Cagnie, G.G.M. Scholten-Peeters
 Writing: R.F. de Best, M.W. Coppieters, E. Van Trijffel, J.C. Bot, R. Castien, G.G.M. Scholten-Peeters
 Data collection: R.F. de Best, A. Compter, J.C. Bot
 Data analysis: R.F. de Best, E. Van Trijffel, G.G.M. Scholten-Peeters
 Project management: J.C. Bot, G.G.M. Scholten-Peeters
 Fund procurement: M.W. Coppieters, G.G.M. Scholten-Peeters
 Providing facilities/equipment: J.C. Bot, G.G.M. Scholten-Peeters
 Providing institutional liaisons: J.C. Bot, G.G.M. Scholten-Peeters
 Clerical/secretarial support: J.C. Bot
 Consultation (including review of manuscript before submitting): M.W. Coppieters, E. Van Trijffel, A. Compter, M. Uyttenbogaart, J.C. Bot, R. Castien, J.J.M. Pool, B. Cagnie, G.G.M. Scholten-Peeters

Acknowledgments

The authors would like to thank all patients and physical therapists who participated in this study.

Funding

This study was funded by a research grant from the Dutch Association for Manual Therapy (Nederlandse Vereniging voor Manuele Therapie) and was supported by the MSG Science Network Physiotherapy (<https://msg-sciencenetwork.nl/>).

Disclosures

The authors completed the ICMJE Form for Disclosure of Potential Conflicts of Interest and reported no conflicts of interest.

References

- Chevan J, Riddle DL. Factors associated with care seeking from physicians, physical therapists, or chiropractors by persons with spinal pain: a population-based study. *J Orthop Sport Phys Ther.* 2011;41:467–476.
- Côté P, Cassidy JD, Carroll L. The treatment of neck and low back pain: who seeks care? who goes where? *Med Care.* 2001;39:956–967.
- Thiel HW, Bolton JE, Docherty S, Portlock JC. Safety of chiropractic manipulation of the cervical spine. *Spine (Phila Pa 1976).* 2007;32:2375–2378.
- Allison GT, Purdam C. Eccentric loading for Achilles tendinopathy—strengthening or stretching? *Br J Sports Med.* 2009;43:276–279.
- Hidalgo B, Hall T, Bossert J, Dugeny A, Cagnie B, Pitance L. The efficacy of manual therapy and exercise for treating non-specific neck pain: a systematic review. *J Back Musculoskelet Rehabil.* 2017;30:1149–1169.
- Coulter ID, Crawford C, Vernon H, et al. Manipulation and mobilization for treating chronic nonspecific neck pain: a systematic review and meta-analysis for an appropriateness panel. *Pain Physician.* 2019;22:E55–E70.
- Falsiroli Maistrolo L, Rafanelli M, Turolla A. Manual therapy and quality of life in people with headache: systematic review and meta-analysis of randomized controlled trials. *Curr Pain Headache Rep.* 2019;23:23–25.
- Lemmens J, De Pauw J, Van Soom T, et al. The effect of aerobic exercise on the number of migraine days, duration and pain intensity in migraine: a systematic literature review and meta-analysis. *J Headache Pain.* 2019;20:16.
- Haynes MJ, Vincent K, Fischhoff C, Bremner AP, Lanlo O, Hankey GJ. Assessing the risk of stroke from neck manipulation: a systematic review. *Int J Clin Pract.* 2012;66:940–947.
- Kranenburg HA, Schmitt MA, Puentedura EJ, Luijckx GJ, van der Schans CP. Corrigendum to “adverse events associated with the use of cervical spine manipulation or mobilization and patient characteristics: a systematic review” [Musculoskelet. Sci. Pract. 28C (2017) 32–38]. *Musculoskelet Sci Pract.* 2018;36:96.
- Thomas LC. Cervical arterial dissection: an overview and implications for manipulative therapy practice. *Man Ther.* 2016;21:2–9.
- Chung CLR, Côté P, Stern P, L’Espérance G. The association between cervical spine manipulation and carotid artery dissection: a systematic review of the literature. *J Manip Physiol Ther.* 2015;38:672–676.
- Debette S. Pathophysiology and risk factors of cervical artery dissection: what have we learnt from large hospital-based cohorts? *Curr Opin Neurol.* 2014;27:20–28.
- Biller J, Sacco RL, Albuquerque FC, et al. Cervical arterial dissections and association with cervical manipulative therapy. *Stroke.* 2014;45:3155–3174.
- Debette S, Leys D. Cervical-artery dissections: predisposing factors, diagnosis, and outcome. *Lancet Neurol.* 2009;8:668–678.
- Cassidy JD, Boyle E, Côté P, Hogg-Johnson S, Bondy SJ, Haldeman S. Risk of carotid stroke after chiropractic care: a population-based case-crossover study. *J Stroke Cerebrovasc Dis.* 2017;26:842–850.
- Nielsen SM, Tarp S, Christensen R, Bliddal H, Klokke L, Henriksen M. The risk associated with spinal manipulation: an overview of reviews. *Syst Rev.* 2017;6:64.
- Rushton A, Rivett D, Carlesso L, Flynn T, Hing W, Kerry R. International framework for examination of the cervical region for potential of cervical arterial dysfunction prior to orthopaedic manual therapy intervention. *Man Ther.* 2014;19:222–228.
- van Trijffel E, Lindeboom R, Bossuyt PM, et al. Indicating spinal joint mobilisations or manipulations in patients with neck or low-back pain: protocol of an inter-examiner reliability study among manual therapists. *Chiropr Man Therap.* 2014;22:22.
- Thoomes-de Graaf M, Thoomes E, Carlesso L, Kerry R, Rushton A. Adverse effects as a consequence of being the subject of orthopaedic manual therapy training, a worldwide retrospective survey. *Musculoskelet Sci Pract.* 2017;29:20–27.
- de Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol.* 2006;59:1033–1039.
- de Vet HCW, Mullender MG, Eekhout I. Specific agreement on ordinal and multiple nominal outcomes can be calculated for more than two raters. *J Clin Epidemiol.* 2018;96:47–53.

23. Kottner J, Audige L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int J Nurs Stud.* 2011;48:661–671.
24. National Institute for Health and Care Excellence. *Hypertension in adults: diagnosis and management.* 2019. Accessed September 5, 2019. <https://www.nice.org.uk/guidance/ng136>.
25. Apeldoorn AT, Den Arend MC, Schuitemaker R, et al. Interrater agreement and reliability of clinical tests for assessment of patients with shoulder pain in primary care. *Physiother Theory Pract.* 2021;37:177–196.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
27. Arnold C, Bourassa R, Langer T, Stoneham G. Doppler studies evaluating the effect of a physical therapy screening protocol on vertebral artery blood flow. *Man Ther.* 2004;9:13–21.
28. Beckers P, Goossens P, Cattrysse E. The reliability of neck positioning during the premanipulative hold. *Ital J Physiother.* 2012;2:135–141.
29. Kaale BR, Krakenes J, Albrektsen G, Wester K. Clinical assessment techniques for detecting ligament and membrane injuries in the upper cervical spine region—a comparison with MRI results. *Man Ther.* 2008;13:397–403.
30. Hutting N, Verhagen AP, Vijverman V, Keesenberg MDM, Dixon G, Scholten-Peeters GGM. Diagnostic accuracy of premanipulative vertebrobasilar insufficiency tests: a systematic review. *Man Ther.* 2013;18:177–182.
31. Osmotherly PG, Rivett DA, Rowe LJ. The anterior shear and distraction tests for craniocervical instability. An evaluation using magnetic resonance imaging. *Man Ther.* 2012;17:416–421.
32. Cassidy JD, Boyle E, Côté P, et al. Risk of vertebrobasilar stroke and chiropractic care. *Spine (Phila Pa 1976).* 2008;33:S176–S183.
33. Vaughan B, Moran R, Tehan P, et al. Manual therapy and cervical artery dysfunction: identification of potential risk factors in clinical encounters. *Int J Osteopath Med.* 2016;21:40–50.
34. Rothwell DM, Bondy SJ, Williams JI. Chiropractic manipulation and stroke. A population-based case-control study. *Stroke.* 2001;32:1054–1060.