UNIVERSITÀ DEGLI STUDI
DI GENOVA

campus savona

Università degli Studi di Genova

Scuola di Scienze Mediche e Farmaceutiche

Dipartimento di Neuroscienze, Riabilitazione, Oftalmologia, Genetica e Scienze Materno-Infantili

Master in Riabilitazione dei Disordini Muscoloscheletrici

A.A 2014/2015

Campus Universitario di Savona

Affidabilità, validità e rilevanza clinica delle tecniche di palpazione e valutazione manuale della mobilità passiva del rachide lombare e dorsale: revisione della letteratura.

Candidato:

Dott. Ft. Chiara Dal Pos

Relatore:

dott. Ft. OMT Aldo Ciuro

INDICE

1. ABSTRACT *A cura di Chiara Dal Pos e Giulia Franco* 1

2. INTRODUZIONE *A cura di Chiara Dal Pos e Giulia Franco* 3

3. MATERIALI E METODI *A cura di Chiara Dal Pos e Giulia Franco* 6

4. RISULTATI *A cura di Chiara Dal Pos e Giulia Franco* 9

4.1. TABELLA SINOTTICA RACHIDE LOMBARE *A cura di Chiara Dal Pos* 11

4.2. TABELLA DEGLI STUDI ESCLUSI PER INSUFFICIENTE QUALITA' METODOLOGICA - RACHIDE LOMBARE *A cura di Chiara Dal Pos* 29

4.3. RISULTATI - RACHIDE LOMBARE *A cura di Chiara Dal Pos* 31

4.3.1. Identificazione dei punti di repere 32

4.3.2. Valutazione della mobilità 33

4.4. TABELLA SINOTTICA RACHIDE TORACICO *A cura di Giulia Franco* 35

4.5. TABELLA DEGLI STUDI ESCLUSI PER INSUFFICIENTE QUALITA' METODOLOGICA – RACHIDE TORACICO *A cura di Giulia Franco* 42

4.6. RISULTATI RACHIDE TORACICO *A cura di Giulia Franco* 43

4.6.1 Identificazione dei punti di repere 44

4.6.2 Valutazione della mobilità 45

5. DISCUSSIONE 48

5.1. RACHIDE LOMBARE *A cura di Chiara Dal Pos* 48

5.2. RACHIDE TORACICO *A cura di Giulia Franco* 53

6. CONCLUSIONI *A cura di Chiara Dal Pos e Giulia Franco* 58

RINGRAZIAMENTI 59

BIBLIOGRAFIA 60

1. ABSTRACT

Disegno di studio. Revisione sistematica.

Obiettivo. Analizzare il livello di evidenza della letteratura corrente con focus sull'affidabilità intra- e inter-esaminatore, sulla validità e sulla rilevanza clinica di esame palpatorio e test di mobilità passiva applicati al rachide dorso-lombare.

Background. Per quanto l'utilizzo della palpazione per identificare i punti di repere ed i test di mobilità passiva sulla base degli stessi siano ampiamente utilizzati dai terapisti manuali nell'esame valutativo del rachide dorso-lombare, la letteratura sull'argomento non sembra presentare concordanti e consistenti evidenze su accuratezza e riproducibilità di tali tecniche.

Materiali e metodi. Gli autori hanno esaminato la letteratura degli ultimi dieci anni utilizzando la banca dati Pubmed. La ricerca, iniziata a settembre 2015 e conclusasi a marzo 2016, è stata effettuata da due revisori indipendenti. Il consenso di un terzo esperto ha eliminato le discrepanze.

Una prima esclusione degli articoli estratti dalla stringa di ricerca è stata creata tramite la pertinenza del titolo. La seconda scrematura è avvenuta con la lettura dell'abstract. Gli articoli inclusi sono stati sottoposti ad analisi del full text, escludendo pertanto quelli di cui non è stato possibile reperire l'intero testo oltre a quelli che ad avvenuta lettura non rispondevano ai criteri di inclusione prestabiliti, e a valutazione della qualità metodologica.

Risultati. Sono stati selezionati 18 e 7 studi di buona qualità riguardanti rispettivamente rachide lombare e dorsale. La palpazione dei reperi è stata trattata da un numero maggiore di studi (4 dorsali e 11 lombari) rispetto alla valutazione della mobilità (3 dorsali e 8 lombari). L'affidabilità è stata maggiormente studiata (5 lavori dorsali e 16 lombari) a fronte di una maggiore difficoltà nel raggiungere una buona qualità metodologica nello studiare la validità (2 lavori dorsali e 5 lombari).

I risultati raccolti sono difficilmente confrontabili a causa delle differenti tecniche o combinazioni di tecniche utilizzate e dei diversi criteri d'inclusione ed obiettivi di studio.

L'affidabilità, con valori che si estendono da un basso range nei test di mobilità del rachide lombare a un moderato livello nella mobilità della regione dorsale, presenta ad ogni modo

migliori risultati rispetto alla validità, generalmente povera, soprattutto quando viene analizzata l'affidabilità intraoperatore. I valori di entrambi i parametri tendono a salire se la valutazione viene effettuata con combinazione di tecniche anziché per mezzo di un singolo approccio.

Si rilevano risultati discordanti nel ricercare una correlazione con esperienza dell'esaminatore, anomalie strutturali, caratteristiche antropometriche.

Conclusioni. Considerata la scarsità di buoni risultati di validità ed affidabilità, nonché la difficoltà di reperire studi di adeguata qualità metodologica, si auspica una direzione di ricerca volta alla produzione di studi di elevato livello qualitativo che prendano in esame un campione di soggetti eterogeneo e vicino alla popolazione reale e osservino una combinazione di approcci valutativi piuttosto che una singola tecnica.

Un'ulteriore attenzione può essere rivolta ad una ricerca, attraverso RTC, finalizzata all'outcome, ossia al cambiamento della prognosi del paziente in base all'applicazione o meno dei suddetti test diagnostici per rilevarne l'importanza in funzione dei risultati di trattamento.

Parole chiave. Motion palpation, mobility, anatomic landmarks, thoracic and lumbar spine, validity, sensitivity, specificity, accuracy, reliability, consistency, reproducibility

2. INTRODUZIONE

L’esame palpatorio per identificare i punti di repere ed i test di mobilità passiva sulla base degli stessi sono tecniche di valutazione di uso comune per i terapisti manuali nell’esame del rachide dorso-lombare. La maggior parte delle patologie, specialmente legate alla regione lombare, ha infatti cause non specifiche e necessita la ricerca di articolazioni disfunzionali, correlate ad ipo/ipermobilità o dolore. Oltre a ciò, la mobilità disfunzionale delle vertebre e delle articolazioni costali è stata associata a varie sindromi dolorose che coinvolgono distretti limitrofi (altre regioni della colonna vertebrale, arti superiori, arti inferiori) con origine muscolo-scheletrica.

Palpazione statica ed esame della mobilità vengono citate da più fonti come strumenti ordinari utilizzati per capire meglio i disordini legati al LBP, valutare interventi, sottoclassificare [1].

Nella valutazione della mobilità segmentale passiva, viene contemporaneamente registrata l'eventuale risposta dolorosa da parte del paziente: questa viene integrata alle altre informazioni della valutazione per definire più precisamente il quadro diagnostico e lo sviluppo del piano di trattamento.

Il 98% dei terapisti ha infatti basato almeno una parte delle loro decisioni di trattamento sui risultati della valutazione della mobilità segmentale [2].

Ad esempio il sistema di classificazione di trattamento proposto da Fritz et al [3] include l'ipomobilità e l'ipermobilità spinale lombare come uno dei criteri per decidere se la manipolazione vertebrale o gli esercizi di stabilizzazione, rispettivamente, sono appropriati.

La stessa valutazione di mobilità prevede però un requisito alla base della corretta valutazione: l’accuratezza nell’identificazione dei segmenti vertebrali.

Diversi sono gli autori, appartenenti a differenti ambiti di formazione, che si sono focalizzati su tale argomento di ricerca.

L’abilità del clinico di localizzare con consistente accuratezza i livelli spinali potrebbe non solo aumentare l’affidabilità palpatoria [4], ma essere prerequisito per il successo di diverse procedure applicate al rachide (blocchi epidurali, esame EMG, tecniche fisioterapiche e manipolazioni) ed è importante anche per la comparazione di rilevazioni diagnostiche e l’efficiente comunicazione tra operatori che esaminano lo stesso paziente, con conseguenti migliorie per la gestione del paziente stesso e l’efficacia del trattamento [5].

La problematica principale che inficia accuratezza ed affidabilità delle suddette procedure è la difficoltà nell'identificare i reperi. La complessità nella localizzazione delle strutture anatomiche può essere inoltre la giustificazione al maggior accordo intraoperatore rispetto all'affidabilità interoperatore [6].

L'errore secondo alcuni autori nasce dal fatto che vi è una mancata standardizzazione delle tecniche palpatorie [7] mentre la scarsa affidabilità può essere dovuta ad inaccurata identificazione dei reperi, varianti anatomiche (sacralizzazione lombare) nella morfologia ossea, costituzione corporea [4].

Molteplici sono infatti i reperi anatomici utilizzati per stabilire il livello vertebrale. Alcuni clinici partono dalla scapola, dalla 10-12^costa, da C7, che però sono lontane da rachide lombare ed aumentano pertanto la possibilità d'errore; spesso viene utilizzata la cresta iliaca, che però presenta povera affidabilità per variabilità anatomica tra soggetti [8]; è stata suggerita da alcuni autori la palpazione della SIPS, più vicina al rachide lombare, corrispondente al margine inferiore di S2, piuttosto che la linea di Tuffier/Jacoby, ossia la linea che unisce le creste iliache e interseca L4 o spazio intervertebrale L4-L5 [4],[9].

Un'ulteriore problematica che sottende alla palpazione è la quantificazione della validità dei diversi test in relazione ad un gold standard applicabile. La palpazione effettuata da esperti in terapia manuale si è dimostrata non essere così accurata da essere usata come reference standard per la validazione di ulteriori approcci e test.

Permane quindi la necessità di confronto con strumenti di imaging quali Rx, RM, ultrasuono, che per ovvi motivi non sono di facile applicabilità e devono comunque prevedere la considerazione di possibili movimenti della pelle durante gli spostamenti dalla posizione della valutazione a quella per l'imaging [1].

La valutazione della mobilità presenta ancor maggiori difficoltà nell'identificazione di un reference standard, tant'è che non vi è quindi ancora chiarezza univoca su un metodo clinico per l'identificazione del livello vertebrale e relativi movimenti accessori che sia accurato, ripetibile, non invasivo e di facile applicabilità.

L'effettivo utilizzo di tali metodiche volte ad identificare reperi ossei e valutare la mobilità passiva del rachide nella generale pratica clinica dei terapisti manuali, rimane svincolato da quelli che

sono i risultati riportati dalla letteratura scientifica. L'abitudine costante di basare infatti l'impostazione delle procedure di trattamento sui parametri rilevati dai test di palpazione e mobilità accessoria non trova supporto nella scarsità di concordanti e consistenti evidenze su accuratezza e riproducibilità delle suddette tecniche.

Alla luce di queste considerazioni, **l'obiettivo dello studio** sarà volto ad analizzare il livello di evidenza della letteratura corrente con focus sull'affidabilità intra- ed inter-esaminatore, sulla validità e sulla rilevanza clinica di esame palpatorio e test di mobilità passiva applicati al rachide dorso-lombare attraverso una revisione sistematica che includerà solamente studi di elevata qualità metodologica; fondamentale infatti è ridurre al minimo il rischio di ottenere dati influenzati da scarsa validità dei risultati in quanto la maggioranza della letteratura in questione presenta studi di carente qualità metodologica.

Nello specifico, sarà attenzione degli autori verificare i test più frequentemente investigati, la tipologia di esaminatori e la scelta della popolazione di pazienti, i possibili bias, l'eventuale correlazione tra risultati e parametri dipendenti dall'operatore (esperienza) o dal soggetto in esame (caratteristiche antropometriche).

3. MATERIALI E METODI

Gli autori hanno esaminato la letteratura degli ultimi dieci anni utilizzando la banca dati Medline attraverso il motore di ricerca dedicato PubMed.

La ricerca, iniziata a settembre 2015 e conclusasi a marzo 2016, è stata effettuata in cieco ed in modo parallelo da due operatori indipendenti aventi come obiettivo lo stesso quesito clinico e col compito specifico postumo di discutere separatamente i risultati relativi al rachide dorsale e al rachide lombare. Il consenso di un terzo esperto ha eliminato le discrepanze.

Una prima esclusione indipendente degli articoli estratti dalla stringa di ricerca è stata determinata tramite la pertinenza del titolo. La seconda scrematura è avvenuta con la lettura dell'abstract.

Gli articoli inclusi a seguito di un confronto tra i tre revisori sono stati sottoposti ad analisi del full text e a valutazione della qualità metodologica, escludendo pertanto quelli di cui non è stato possibile reperire l'intero testo oltre a quelli che ad avvenuta lettura non rispondevano ai criteri di inclusione prestabiliti.

La stringa di ricerca ottenuta dal confronto tra i revisori è pertanto la seguente:

> (((((palpat* OR mobili* OR anatomic landmark OR anatomic landmarks OR accessory movement)) AND (spine OR thoracic OR lumbar OR vertebra*)) AND (reliability OR accuracy OR consistency OR reproducibility OR validity OR sensitivity OR specificity)) AND "last 10 years"[PDat] AND English[lang]) Sort by: Title Filters: published in the last 10 years; English.

Criteri di inclusione: sono stati considerati oggetto di valutazione gli articoli di ricerca primaria aventi come argomento l'affidabilità intra-extraoperatore, la validità, la rilevanza clinica dell'esame palpatorio e dei test di mobilità e come distretto in oggetto il rachide dorso-lombare, che abbiano ottenuto un adeguato punteggio alla valutazione metodologica, pubblicati o accettati per pubblicazione negli ultimi dieci anni ed in lingua inglese.

L'analisi metodologica è stata eseguita sia per gli studi di validità che per quelli di affidabilità interoperatore ed intraoperatore. Sono stati utilizzati i criteri espressi dalla checklist di Hestbœk [10] per validità (Fig. 1) e Haneline [11] per affidabilità (Fig. 2), differenziando gli articoli di buona qualità da quelli di scarsa qualità in base al cut off da essi impiegato: studi sull'affidabilità

interesaminatore che hanno raggiunto 3 o più punti (≥50%) sono stati considerati di alta qualità, mentre quelli che hanno ottenuto meno di tre punti (<50%) sono stati ritenuti di bassa qualità. Per l'affidabilità intraoperatore il cut off era di 2 punti. Studi di validità che hanno raggiunto punteggi ≥80% vengono accettati come di buona qualità.

Criteri di esclusione: sono stati esclusi dallo studio tutti gli articoli che non rispondessero ai criteri di inclusione, che non avessero come argomento la terapia manuale, che non riportassero dati separati e differenziati per ogni test clinico in esame, senza test statistici adeguati; test clinici che non rispondessero agli outcomes prefissati (rilevazione dei punti di repere e valutazione della mobilità passiva) o utilizzo di test di palpazione per valutare altri outcomes primari (identificazione del centro di rotazione di articolazioni specifiche, ROM); articoli con obiettivo di diagnosticare patologie del rachide di tipo strutturale (fratture, scoliosi, spondilolistesi, protrusioni, ernie discali) e/o funzionale oppure che ne studiassero approccio e trattamento adatti; articoli che valutassero l'affidabilità dei test di mobilità in disordini muscolo-scheletrici diversi da lombalgia o dorsalgia aspecifica.

Figura 1. Checklist per la valutazione metodologica della validità. (Hestbœk, 2000)

FOR VALIDITY STUDIES:

1. 1.1. Test validated against "golden measure" 3
   - 1.2. Or symptoms or previously validated tests 2
   - 1.3. Or nonvalidated tests or treatment/no treatment 1

   (maximum possible for validation method is 3)
2. Examiner blind to all applicable aspects 1
3. Subject naïve/without vested interest/unable to affect result 1
4. Description of test procedure and validation method 1
5. Study subjects and examiners described 1
6. Description of statistics and significance level defined 1
7. Clear description of protocol 1
8. No additional problems 1

Maximum quality score 10

Figura 2. Checklist per la valutazione metodologica dell'affidabilità. (Haneline M.T., 2009)

SIX-POINT SCALE USED TO ASSESS THE QUALITY OF STUDIES. QUESTIONS 2, 4, 5, AND 6 APPLY TO INTRAEXAMINER STUDIES

1. Order of observers conducting the test(s) randomized (1 point)
2. Case mix (1 point) (Both symptomatic and asymptomatic subjects. For studies with an intentional case mix, the case mix must represent a natural clinic population, that is, a population consisting of a variety of patients with regards to sex, age, and problems in different part of the spine. If the purpose is to look at subgroups, such as asymptomatic patients only, credit will be given.)
3. Observers blind to other observers findings (1 point)
4. Observers blind to confounding information (1 point) (such as patient history)
5. Subjects blind to observer findings (1 point for true/complete blinding, 0.5 point for no blinding but subjects unable to affect results)
6. Kappa (κ, κo, κg, κw) or ICC (l-way ICC, 2-way ICC, or generalizability coefficients for single rating by each observer) used for analysis (1 point)

Interexaminer studies that scored 3 or more points (≥50%) were considered to be high quality, and those scoring less than 3 (b50%) low quality.
Intraexaminer studies that scored 2 or more points (≥50%) were considered to be high quality, and those scoring less than 2 (b50%) low quality.

4. RISULTATI

La stringa di ricerca ha prodotto 722 risultati. Sono stati selezionati per titolo 41 articoli di cui è stata effettuata la lettura dell'abstract, in cieco dai due autori, con una successiva supervisione del terzo revisore.

Il confronto tra le parti ha visto un accordo unanime su 21 articoli; 28 articoli sono stati selezionati da almeno due revisori ed altri 7 sono stati presi in esame da uno di essi.

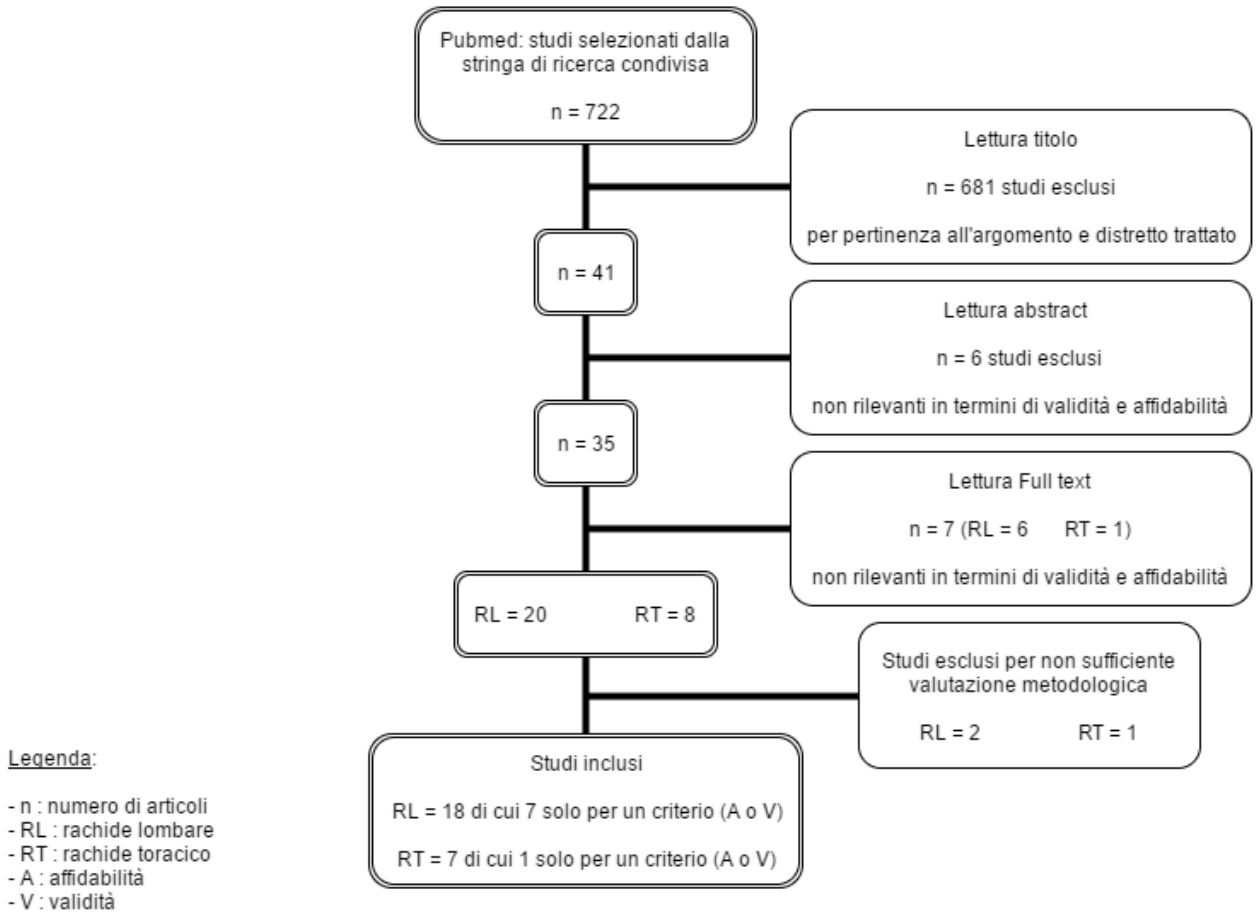
Sono stati pertanto selezionati in base all'abstract 25 articoli relativi al rachide lombare, 8 specifici per il rachide dorsale e 2 riguardanti entrambi i distretti, la cui lettura del full test e conseguente valutazione metodologica è stata eseguita in modo indipendente da ciascuno dei due revisori in base al proprio argomento di competenza.

I criteri di esclusione per titolo e per abstract si sono basati sulla tipologia di letteratura e sulla pertinenza dell'argomento trattato e del distretto valutato.

Alla lettura del full test sono stati ulteriormente esclusi nove articoli riguardanti il rachide lombare e due relativi al rachide dorsale perché non rilevanti o insufficienti alla valutazione metodologica (Tabelle 4.2 e 4.5) di validità ed affidabilità; sono stati dunque utilizzati nella loro interezza rispettivamente undici articoli d'argomento lombare e sei articoli per il rachide dorsale; sette articoli di pertinenza lombare ed uno dorsale sono stati considerati solo per l'aspetto di validità o affidabilità risultato superiore al cut off di qualità metodologica, per un totale di diciotto studi lombari e sette lavori dorsali considerati per la stesura dei risultati di revisione (Fig. 3).

La qualità metodologica degli studi analizzati non si è rivelata dipendente dagli anni di pubblicazione.

Figura 3. Flow chart della selezione degli articoli inseriti nella revisione



Nelle tabelle sinottiche seguenti vengono riportate schematicamente le caratteristiche più rilevanti di ogni studio ed i riferimenti degli studi esclusi perché insufficienti alla valutazione metodologica, accompagnate dal capitolo descrittivo relativo a ciascun distretto.

4.1. TABELLA SINOTTICA RACHIDE LOMBARE

<table>
<tr><td colspan="7">1. Interexaminer reliability of lumbar segmental mobility tests. (F. Johansson, 2005) [12]</td></tr>
<tr><td>Distretto</td><td colspan="2">Obiettivo</td><td>Esaminatori</td><td>Soggetti</td><td>Gold standard (validità)</td><td>Risultati</td></tr>
<tr><td>Rachide lombare</td><td colspan="2">Verificare l’affidabilità interesaminatore sulla valutazione della mobilità segmentaria del rachide in estensione e in flessione con test di mobilità passiva.</td><td>3 terapisti OMT di 2° livello che non hanno mai lavorato assieme.</td><td>30 soggetti tra pazienti con LBP pregresso e studenti asintomatici.</td><td></td><td>Scala di valori in 5 punti.
K pesato <0.20 in 24 dei 30 soggetti.
K pesato rilevato “moderato” in L2-L3 e L5-S1 in estensione tra ft 2 e ft 3.
Portando la scala di valori da 5 punti a 3 punti il k pesato non differiva in maniera sostanziale.
Povera affidabilità inter-ex tra clinici OMT che non lavoravano insieme e non avevano possibilità di standardizzare la tecnica insegnata.</td></tr>
<tr><td colspan="7">AFFIDABILITA’ INTEROPERATORE: 5 (83%)</td></tr>
<tr><td colspan="2">Order of observer randomized: 1</td><td>Case mix: 1</td><td>Observer blind to other observers findings: 1</td><td>Observer blind to counfounding informations: 0. Ft conoscono storia clinica.</td><td colspan="2">Subjects blind to observer findings: 1</td><td>K/ICC: 1. K e K pesato</td></tr>
</table>

2. Identification of the correct lumbar level using passive intersegmental motion testing. (R. C. Chakraverty, 2007) [13]							
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**		
Rachide lombare	Valutazione di accuratezza e concordanza interoperatore con il test di mobilità passiva intersegmentale per l'identificazione dello spazio intervertebrale L5-S1, tra operatore esperto e novizio dopo un'ora di training (Gruppo 1) e dopo tre mesi (Gruppo 2).	Un medico specializzato in rachide e con formazione osteopatica, uno studente al 4° anno di medicina.	35 pazienti reclutati per infiltrazioni guidate tramite fluoroscopia per LBP.	Fluoroscopia	Esperto: accuratezza del 57% con miglioramento significativo tra Gruppo 1 2 gruppo 2 (31%-79%), a dimostrazione di capacità di apprendimento. Inesperto: accuratezza media del 54% con incremento 38%-68% non statisticamente significativo. Concordanza povera. Accordo interesaminatore: j = 0.26 (95% CI: da –0.03 a 0.53), j = 0.01 (95% CI: da 0.31 a 0.40), e j = 0.20 (95% CI: da –0.02 a 0.44) rispettivamente per Gruppo 1, 2, ed entrambi. Non relazione tra altezza-peso-BMI-età e identificazione.		
AFFIDABILITA' INTEROPERATORE: Non sufficiente							
VALIDITA': 8							
Test validated against gold standard: 2	*Examiner blind to all applicable aspects: 0,5*	*Subject naive/without vested interest/unable affected result: 1*	*Description of test procedure and validation method: 1*	*Study subjects and examiners described: 1*	*Descriptions of statistics an significance level defined: 1*	*Clear description protocol: 0,5*	*No additional problems: 0,5*

3. Palpation identification of spinous processes in the lumbar spine. (J.C. Harlick, 2007) [7]							
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**		
Processi spinosi lombari	Verificare accuratezza nella palpazione, con tecnica a scelta, dei processi spinosi lombari (L1-L3-L5.	5 terapisti esperti (10 anni di clinica) e formati in terapia manuale.	75 pazienti con prescrizione non urgente RX per LBP (ogni fisioterapista ne valuta 15).	RX antero-posteriore e laterale.	72% accuratezza per livello esatto o adiacente; 47% (tra 24 e 69%) accuratezza per esatto livello. Variazione statisticamente significativa dell'accordo tra terapisti in L3 (OR=1.56, CI=1.10–2.23, P=0.01) e L5 (OR=1.74, CI=1.21–2.51, P<0.01). Età, sesso, BMI non correlate (P>0.05). L1 presenta un minor spessore dei tessuti molli, che diventa progressivamente maggiore fino ad L5. Il minor spessore nei tessuti molli ha dimostrato debole correlazione con accuratezza solo in L5 (P=0.05). Effetto predittivo dell'altezza del processo spinoso per l'accuratezza della palpazione: debole per L5 (P=0.04), più forte per L3 (P=0.01). La scelta della tecnica palpatoria o associazione non influenza la previsione dell'accuratezza (P>0.05).		
VALIDITA': 8,5							
Test validated against gold standard: 3	*Examiner blind to all applicable aspects: 0,5*	*Subject naive/without vested interest/unable affected result: 1*	*Description of test procedure and validation method: 1*	*Study subjects and examiners described: 1*	*Descriptions of statistics and significance level defined: 1*	*Clear description protocol: 0,5*	*No additional problems: 0,5*

4. Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing. (R. Landel, 2007) [14]							
Distretto	**Obiettivo**		**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**	
Rachide lombare	Verificare l'affidabilità interesaminatore nella valutazione in PA della mobilità segmentaria nel rachide lombare e la validità con MRI.		2 fisioterapisti con 15 (PA con MRI) e 16 (PA senza MRI) anni di esperienza in terapia manuale. Terzo operatore registra dati MRI.	29 soggetti tra 18 e 45 anni con diagnosi di LBP aspecifico.	Risonanza Magnetica dinamica	Buona affidabilità nell'identificare il segmento meno mobile (accordo=82.8%, kappa=71,95%, inter-vallo di confidenza [CI]=da .48 a .94). Povera affidabilità nell'identificare il segmento più mobile (Kappa=29,95%, CI=da -.13 a .71) nonostante il buon accordo (79,3%). Povero accordo tra valori PA e movimento intervertebrale misurato con MRI (Kappa=04,95% , CI=da -.16 a .24 per il segmento meno mobile, Kappa=00,95% CI=da -.09 a .08, per il segmento più mobile). Affidabilità intraoperatore, misurata con MRI su 5 soggetti, rilevata eccellente. ICC da .95 a .99.	
AFFIDABILITA' INTEROPERATORE: 3 (50%)							
Order of observer randomized: 0	*Case mix: 0. Fascia d'età ristretta, gruppo con LBP.*		*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 0*	*Subjects blind to observer findings: 1*	*K/ICC: 1. K e ICC.*	
VALIDITA': 8							
Test validated against gold standard: 3	*Examiner blind to all applicable aspects: 0,5*	*Subject naive/without vested interest/unable affected result: 0,5*	*Description of test procedure and validation method: 1*	*Study subjects and examiners described: 1*	*Descriptions of statistics and significance level defined: 1*	*Clear description protocol: 0,5*	*No additional problems: 0,5*

5. Interexaminer reliability and accuracy of posterior superior iliac spine and iliac crest palpation for spinal level estimations. (H.W. Kim, 2007) [9]								
Distretto	**Obiettivo**		**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**		
SIPS e cresta iliaca	Verificare l'affidabilità interoperatore nella palpazione di SIPS e linea di Tuffier, poi esaminare l'accuratezza attraverso RX.		4 esaminatori con esperienza almeno triennale in medicina muscoloscheletrica. 4 esaminatori per raccolta dati.	60 pazienti per affidabilità interesaminatore; 72 pazienti per accuratezza.	Rx postero-anteriori per accuratezza.	Affidabilità interoperatore significativamente maggiore per SIPS rispetto a cresta iliaca (K=0.92/0.96 per SISP, K=0.60/0.59 per cresta iliaca, P<0.05). Livelli spinali corrispondenti a SIPS: da spazio intervertebrale L5-S1 a spinosa S2, con S1 quale livello più frequentemente rilevato (31 [43%] di 72 casi); Livello stimato della cresta iliaca (correttamente identificato nel 36% dei casi, ossia 26): da L2-L3 a spinosa L5. Linea di Tuffier localizzata tra il processo spinoso di L4 e l'interspazio L5-S1. Maggior difficoltà nel comparare accuratezza dell'identificazione del livello spinale tramite la SIPS rispetto alla CI.		
AFFIDABILITA' INTEROPERATORE: 3(50%)								
Order of observer randomized: 0. Non specificato.	*Case mix: 0*		*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 0*		*K/ICC: 1 X^2 test, p value.*	
VALIDITA': 8,5								
Test validated against gold standard: 3	*Examiner blind to all applicable aspects: 1*	*Subject naive/without vested interest/unable affected result: 1*	*Description of test procedure and validation method: 0,5*	*Study subjects and examiners described: 1*	*Descriptions of statistics and significance level defined: 1*	*Clear description protocol: 0,5*	*No additional problems: 0,5*	

6. Reliability of lumbar spinal palpation, range of motion, and determination of position. (M. Troke, 2007) [15]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Rachide lombare	Esaminare il livello di accordo interesaminatore nell'esame palpatorio dei processi spinosi lombari, la ripetibilità test-retest intraoperatore del ROM e l'affidabilità delle misure in stazione eretta in soggetti asintomatici.	Tre operatori. Competenze non specificate.	22 pazienti asintomatici, dai 22 ai 38 anni.		19 su 22 palpazioni accurate (86,4%). Nella stazione eretta, i valori ICC sul piano sagittale, coronale e orizzontale erano rispettivamente 0.96, 0.80, e 0.98. Nonostante le differenze statisticamente significative rilevate dagli esaminatori nelle diverse posizioni di flesso-estensione, lateroflessione e rotazione ($p < 0.05$), l'entità delle differenze (2 gradi o meno) non è stata ritenuta clinicamente importante.
AFFIDABILITA' INTEROPERATORE:4 (67%), AFFIDABILITA' INTRAOPERATORE: 3 (75%)					
Order of observer randomized: 1	*Case mix: 0*	*Observer blind to other observers findings: 0. Markers sui reperi visibili.*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 1*	*K/ICC: 1. ICC*

7. Spinal palpation for lumbar segmental mobility and pain provocation: an interexaminer reliability study. (Schneider, 2008) [16]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Rachide lombare e articolazione sacro-iliaca	Verificare l'affidabilità interesaminatore di: PA per provocazione dolore; PA per mobilità segmentale; Prone instability test.	Due esaminatori chiropratici con 25 e 10 anni di esperienza, con due ore di training per concordare protocollo e procedure d'esame.	39 pazienti con LBP tra i 18 e i 65 anni.		Dati suddivisi in Lombare superiore e Lombare inferiore. Affidabilità per mobilità segmentaria povera: kappa da −.17 a .17. Pain provocation test da discreto a buono: kappa da .21 a .73) Prone instability test ragionevole: κappa da .46 a .54.
AFFIDABILITA' INTEROPERATORE: 3 (50%)					
Order of observer randomized: 0. (pz invertiti di volta in volta)	*Case mix: 0. Sottopopolazione con LBP.*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 0*	*K/ICC: 1. (unadjusted k – adjusted k)*

8. Reliability and validity of a palpation technique for identifying the spinous processes of C7 and L5. (R. Robinson, 2009) [5]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Processi spinosi di C7 e L5	Esaminare l'affidabilità interoperatore e la validità di tecniche palpatorie per C7 e L5, con RX di controllo, ricercando l'eventuale correlazione con BMI e sesso. Palpazione effettuata nella stessa posizione dell'esame radiologico.	Due terapisti con 16-18 anni di esperienza post diploma IFOMPT. Radiologo con 30 anni di esperienza.	31 pazienti tra i 26 e i 79 anni, 21 F e 10 M, con prescrizione di RX.	RX in proiezione AP e laterale	Affidabilità intertester povera per C7 e moderata per L5. Accordo totale interoperatore: 11 (35%); accordo<10mm: 22 (71%); accordo <20mm 27 (87%). Processo spinoso di L5 correttamente identificato dai due TM in 15 (48%) e 11 (36%) partecipanti, rispettivamente. Accordo tra i due in 9 (29%) partecipanti. Per 14 (45%) partecipanti, entrambi I TM hanno indicato un punto esterno al target. Kappa 0.48 (95% CI: 0.18–0.78). Nè peso nè sesso influenzano l'accuratezza della palpazione. Alta affidabilità del test palpatorio nell'identificare il segmento target.
AFFIDABILITA' INTEROPERATORE: 3 (50%)					
Order of observer randomized: 1	*Case mix: 1*	*Observer blind to other observers findings: 0*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 0*	*K/ICC: 0*
VALIDITA': Non sufficiente					

9. Allowing a possible margin of error when assessing student skills in spinous process location. (Hart, 2011) [17]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
C2, T3, T9, L2	Determinare l'affidabilità interoperatore e definire il margine d'errore nella palpazione dei processi spinosi di C2, T3, T9, L2.	Due chiropratici ricercatori.	18 studenti.		I valori delle differenze interesaminatori si sono dimostrati in crescita in senso cranio-caudale. C2, 4.23+3.77mm; T3, 13.41+10.53mm; T9, 18.17+17.62mm; L2, 18.70+16.58 mm. L2: ICC 0.910, p-value .000 Tale differenza comprensiva della Deviazione Standard può essere considerata come margine di errore lecito permesso nella valutazione degli studenti.
AFFIDABILITA' INTEROPERATORE: 5 (83%)					
Order of observer randomized: 1. Variati ma non randomizzati.	*Case mix: 0*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 1*	*K/ICC: 1*

10. Palpatory accuracy of lumbar spinous processes using multiple bony landmarks. (K. T. Snider, 2011) [18]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Rachide lombare	Esaminare accuratezza e affidabilità interoperatore nell'identificare i processi spinosi dal L1 a L4, usando una combinazione di diversi reperi ossei (base del sacro, L5, linea di Tuffier, T12, 12° costa), nonché valutare la correlazione con esperienza dell'esaminatore, presenza di anomalie anatomiche, caratteristiche del partecipante.	Tre ricercatori specializzati in medicina neuromuscolare e medicina manipolativa osteopatica, uno studente al terzo anno della medesima specializzazione.	60 soggetti dai 20 ai 60 anni con diverse caratteristiche strutturali e livelli di attività fisica.	RX postero-anteriori	Corretta identificazione nel 69%.(κ = 0.81; CI 95%=0.79-0.83). Diminuzione d'accuratezza statisticamente significativa con obesità (P = .0003) per L3 (50% vs 73%) ed L4 (44% vs 72%), così come per anomalie 12° costa (P≤0.05), con valori in riduzione dal 74% al 55%. Accuratezza maggiore nei maschi rispetto alle femmine (P = .01). I clinici presentano maggior accuratezza dello studente (67%-78% vs 51%, P≤.03). L'accuratezza dipende dall'esperienza dell'esaminatore, dalla presenza di anomalie anatomiche, dalle caratteristiche del partecipante. L'affidabilità interesaminatore tra tutti i partecipanti era moderata, con 62% di accordo e κ = 0.53 (95% CI, 0.49-0.56). L'affidabilità interoperatore tra i clinici era buona con il 73% di accordo e κ = 0.66 (95% CI, 0.61-0.71).
AFFIDABILITA' INTEROPERATORE: 5 (83%)					
Order of observer randomized: 0	*Case mix: 1*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 1*	*K/ICC: 1 (k pesato)*

VALIDITA': 8							
Test validated against gold standard: 3	*Examiner blind to all applicable aspects: 0,5*	*Subject naive/without vested interest/unable affected result: 1*	*Description of test procedure and validation method: 0,5*	*Study subjects and examiners described: 1*	*Descriptions of statistics and significance level defined: 1*	*Clear description protocol: 0,5*	*No additional problems: 0,5*

11. The inter-rater and intra-rater reliability of passive physiological accessory movement assessment of lumbar spine in novice manual therapists. (M. Deore, 2012) [19]

Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Rachide lombare	Verificare affidabilità intra e interoperatore di movimenti accessori fisiologici (tramite PA da prono) passivi al rachide lombare nei fisioterapisti novizi.	Due fisioterapisti formati in terapia manuale e con meno di 12 mesi di esperienza clinica.	52 soggetti sani tra 18-40 anni, senza precedenti esperienze di LBP.		Scala di valori in 3 punti (ipo-normo-iper). Percentuale d'accordo interoperatore varia tra 38% (L4) e 58% (L2). Kappa tra 0.00 (L4) e 0.20 (L2). Percentuale d'accordo intraoperatore tra 38.4% and 57.6%. Kappa tra 0.24 (L2) e 0.57 (L1). Povera affidabilità intra e inter. (accordo accettabile 0.6).

AFFIDABILITA' INTEROPERATORE: 3 (50%), AFFIDABILITA' INTEROPERATORE: 2 (50%)					
Order of observer randomized: 0	*Case mix: 0. Sottopopolazione di soggetti sani.*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 0. Non specificato ma pazienti sani.*	*Subjects blind to observer findings: 1*	*K/ICC: 1 (K)*

<table>
<tr><td colspan="7">12. Manual palpation of lumbo-pelvic landmarks: a validity study. (J. Kilby, 2012) [20]</td></tr>
<tr><td>Distretto</td><td colspan="2">Obiettivo</td><td>Esaminatori</td><td>Soggetti</td><td>Gold standard</td><td>Risultati</td></tr>
<tr><td>Processo spinoso di L4 e SIPS</td><td colspan="2">Verificare l’accuratezza della palpazione manuale del processo spinoso di L4 e di SIPS dx e sx, con controllo ecografico.</td><td>9 fisioterapisti con Master in Terapia Manuale.</td><td>3 soggetti sani.
2M, 1 F; BMI 25,3.</td><td>Ultra-suono</td><td>L4 15.63mm (3.89),
PSIS sx 20.07mm (4.60),
PSIS dx 20.59mm (2.79).
ICC generale 0.81 (p =0.00),
ICC L4 0.83 (p =0.00),
ICC PSIS sx 0.68 (p =0.03),
ICC PSIS dx 0.58 (p =0.06).
Soggetti testati in un’unica sessione, senza cambio di posizione tra test e verifica US.</td></tr>
<tr><td colspan="7">AFFIDABILITA’ INTEROPERATORE: 3 (50%)</td></tr>
<tr><td colspan="2">Order of observer randomized: 0</td><td>Case mix: 0</td><td>Observer blind to other observers findings: 1</td><td>Observer blind to counfounding informations: 0</td><td colspan="2">Subjects blind to observer findings: 1</td><td>K/ICC: 1 (ICC).</td></tr>
<tr><td colspan="7">VALIDITA’: Non sufficiente</td></tr>
</table>

13. Establishing the content validity of palpatory examination for the assessment of the lumbar spine using ultrasonography: a pilot study. (K.Shaw, 2012) [21]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Vertebre lombari	Esaminare la validità di contenuto della valutazione di mobilità segmentale e gli effetti delle HVLAT nelle vertebre disfunzionali lombari, con l'uso dell'ultrasuono.	Un fisioterapista specializzato in medicina manipolativa muscoloscheletrica e osteopatica ed un tecnico radiologo specializzato in ecografia MSK, cieco rispetto alla diagnosi specifica.	12 soggetti (5 M, 7 F), tra i 22 e i 30 anni, sani.	Ecografia	Affidabilità test-retest 0.997 (P<.001).
AFFIDABILITA' INTRAOPERATORE: 3 (75%)					
Order of observer randomized: --	*Case mix: 0. Soggetti asintomatici, escluse anomalie LS.*	*Observer blind to other observers findings: --*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 1*	*K/ICC: 1. Pearson, p value.*
VALIDITA': Non sufficiente					

14. The use of pMRI to validate the identification of palpated bony landmarks. (K. Cooper, 2013) [1]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Processo spinoso di L4	Verificare l'accuratezza della palpazione dei processi spinosi lombari (L4) in stazione eretta e da prono verificando con Risonanza Magnetica statica, nonché l'affidabilità interoperatore.	Due fisioterapisti con almeno 15 anni di esperienza in terapia manuale.	9 partecipanti asintomatici, tra i 18 e i 65 anni, 2 M e 7 F, 6 normopeso e 3 sovrappeso, con assenza di anomalie strutturali al rachide.	pMRI. Palpazione dalla stazione eretta e successiva imaging, palpazione da prono sul lettino dello scan.	71% di markers corretti in posizione prona, 100% di markers corretti in stazione eretta. Accordo interoperatore: 44% in stazione eretta, 38% in posizione prona, che arriva a 89% e 88% per intervallo inferiore a 10mm. pMRI utile per verificare accuratezza palpazione in quanto minimizza gli errori dovuti allo spostamento della cute nei passaggi dalla posizione di palpazione a quella per l'imaging.
AFFIDABILITA' INTEROPERATORE: 3 (50%)					
Order of observer randomized: 1. (con lancio moneta).	*Case mix: 0. Soggetti asintomatici, escluse anomalie LS.*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 0*	*Subjects blind to observer findings: 1*	*K/ICC: 0. Wilcoxon*
VALIDITA': Non sufficiente					

15. Determination of thoracic and lumbar spinal processes by their percentage position between C7 and the PSIS level. (M.J. Ernst, 2013) [22]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
T4, T7, T10, L1, L4	Verificare l'affidabilità interoperatore nel rilevare la posizione dei processi spinosi di T4, T7, T10, L1, L4 come valore percentuale della distanza tra C7 e la SIPS, utilizzando un curvilinee metrico flessibile.	6 fisioterapisti o professionisti affini specializzati in kinesiologia, che effettuavano la misurazione a coppie.	22 soggetti sani tra 22 e 51 anni senza evidenti deviazioni morfologiche del rachide.		Affidabilità relativa pressoché perfetta: ICC >0.9. Affidabilità assoluta: alto accordo interoperatore. L1 rilevata al 70.9% ed L4 all'86.1% della distanza tra C7 e la SIPS. CI massimo ad L1, con intervallo di 2,8% tra il valore massimo e minimo rilevati nei pazienti.
AFFIDABILITA' INTEROPERATORE: 3 (50%)					
Order of observer randomized: 0. Non specificato.	*Case mix: 0. Sani.*	*Observer blind to other observers findings: 1 (sulla misurazione).*	*Observer blind to counfounding informations: 0*	*Subjects blind to observer findings: 1*	*K/ICC: 1 ICC*

16. Validity of palpation techniques for the identification of the spinous process L5. (O. Merz, 2013) [6]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
L5	Verificare l'accuratezza della palpazione di L5 attraverso la motion palpation (MP), l'utilizzo di reperi quali SIPS (PSIS) e cresta iliaca (CI), la motion palpation con i reperi di SIPS e CI visibili (MP+). Valutare l'affidabilità intraoperatore (misurazioni dopo 4 settimane) e interoperatore (fisioterapista e radiologo).	Un terapista OMT, un radiologo.	66 soggetti con LBP.	RX laterali	Affidabilità interoperatore: k=0.637. Affidabilità intraoperatore: k= 0.623. L5 identificata nell'87% da almeno una tecnica, nel 74% da due approcci in accordo. Accuratezza di ogni singola tecnica tra 45% e 61%, senza differenze statisticamente significative tra ciascuna. Se accordo tra due tecniche, l'accuratezza sale dal 58% al 78%. Se accordo tra tre tecniche, l'accuratezza sale dal 69% all'83%. Combinando quindi più tecniche l'accuratezza aumenta. Accordo più frequente tra PSIS e MP+ (78%), accordo meno frequente tra CI e MP (58%). MP più accurata se eseguita come terza tecnica anziché come prima, senza differenze statisticamente significative. No influenze di BMI, genere, età, altezza sull'accuratezza.
AFFIDABILITA' INTRAOPERATORE: 2 (50%); AFFIDABILITA' INTEROPERATORE: Non sufficiente					
Order of observer randomized:	*Case mix: 0. Sottopopolazione di soggetti con LBP ma gruppo eterogeneo per età-sesso-peso-altezza.*	*Observer blind to other observers findings:*	*Observer blind to counfounding informations: 0*	*Subjects blind to observer findings: 1*	*K/ICC: 1. K, IC*
VALIDITA': Non sufficiente					

17. Intertester agreement and validity of identifying lumbar pain provocative movement patterns using active and passive accessory movement tests. (Hidalgo, 2014) [23]

Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
Rachide lombare	Valutare l'affidabilità interoperatore e la validità dei test attivi (flessione attiva singola e ripetuta 10 volte-flessione con sovrapressione, flessione con lateroflessione; estensione come sopra) per identificare una direzione provocativa associata a movimenti accessori passivi di provocazione del dolore (PA in neutra, flessione ed estensione, pressione laterale in neutra).	Fisioterapisti OMT, uno con 10 anni di esperienza, l'altro senza esperienza clinica, in cieco. Terzo supervisore per raccolta dati.	18 soggetti asintomatici e 18 soggetti con LBP subacuto, tra i 20 e i 65 anni, BMI inferiore a 30 kg m^{-2} .	No reference standard. Accuratezza valutata con analisi sensibilità e specificità. Limitazione per validità interna, ma facilmente applicabile a contesto esterno.	Accordo interoperatore da buono a eccellente nell'identificare il pattern doloroso in flessione (κ = 0.87-1) o estensione(κ = 0.65-0.74). Accordo da moderato ad eccellente nell'identificare la dolorabilità coi movimenti accessori intervertebrali (κ = 0.42-0.83). Buona sensibilità (0.67-0.87) e specificità (0.82-0.85) nel distinguere soggetti con LBP da asintomatici usando la combinazione delle due procedure d'esame. Differenze statisticamente significative nella risposta alla provocazione passiva a seconda del livello: P< 0.001 tra L5 e L1-2, L4 ed L1, ed L3 ed L1 nel gruppo LBP; P<0.001 tra L5 ed L1-4 nel gruppo asintomatico.
AFFIDABILITA' INTEROPERATORE: 5 (83%)					
Order of observer randomized: 1	*Case mix: 1*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations:*	*Subjects blind to observer findings: 0*	*K/ICC: 1 K e PABAK*
VALIDITA': Non sufficiente					

18. Inter- and intra-observer agreement of the motion palpation test for lumbar vertebral rotational asymmetry. (M.A. Bracht, 2015) [24]					
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard**	**Risultati**
L2	Verificare l'affidabilità inter e intra-operatore della valutazione in PA dell'asimmetria nella rotazione vertebrale.	3 esaminatori esperti in terapia manuale.	51 soggetti asintomatici.		Da 0,07 a 0,37 intraoperatore; da 0,12 a 0,30 interoperatore. Bassa affidabilità, significato clinico discutibile.
AFFIDABILITA' INTEROPERATORE: 5 (83%), AFFIDABILITA' INTRAOPERATORE : 3 (75%)					
Order of observer randomized: 1	*Case mix: 0. Sottopopolazione di soggetti sani.*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 1*	*K/ICC: 1 (K e K pesato)*

4.2. TABELLA DEGLI STUDI ESCLUSI PER INSUFFICIENTE QUALITA' METODOLOGICA - RACHIDE LOMBARE

Titolo	**Autore, anno, rivista**	**Target (argomento e distretto)**	**Punteggio qualità**
Simple anatomical information improves the accuracy of locating specific spinous processes during manual examination of the low back.	Philips, 2009 *Manual Therapy [25]*	Validità RL (L3) e RT (T12)	Validità 6/10
Soft tissue depression at the iliac crest prominence: a new landmark for identifying the L4-L5 interspace.	Borghi, 2012 *Minerva anestesiologica [8]*	Validità LR (spazio intervertebrale L4-L5)	Validità 7,5/10
Identification of the correct lumbar level using passive intersegmental motion testing.	Chakraverty, 2007 *Anaesthesia [13]*	Affidabilità interoperatore e validità RL (spazio intervertebrale L5-S1)	Affidabilità 33% Validità sufficiente
Reliability and validity of a palpation technique for identifying the spinous processes of C7 and L5.	Robinson, 2009 *Manual Therapy [5]*	Validità e affidabilità interoperatore RC (C7) e RL (L5)	Validità 7,5/10 Affidabilità sufficiente
Manual palpation of lumbo-pelvic landmarks: a validity study	Kilby, 2012 *Manual Therapy [20]*	Validità e affidabilità interoperatore RL (L4) e SIPS	Validità 7,5/10 Affidabilità sufficiente
Establishing the content validity of palpatory examination for the assessment of the lumbar spine using ultrasonography: a pilot study.	Shaw, 2012 *The Journal of the American Osteopathic Association [21]*	Validità e affidabilità intraoperatore RL	Validità 3,5/10 Affidabilità sufficiente

The use of pMRI to validate the identification of palpated bony landmarks.	Cooper, 2013 *Manual Therapy [1]*	Validità a affidabilità interoperatore RL (L4)	Validità 7,5/10 Affidabilità sufficiente
Validity of palpation techniques for the identification of the spinous process L5.	Merz, 2013 *Manual Therapy [6]*	Validità, affidabilità interoperatore e intraoperatore	Validità 7,5/10 Affidabilità interoperatore 33% Affidabilità intraoperatore sufficiente
Intertester agreement and validity of identifying lumbar pain provocative movement patterns using active and passive accessory movement tests.	Hidalgo, 2014 *Journal of Manipulative and Physiological Therapeutics [23]*	Validità e affidabilità interoperatore RL	Validità 6,5/10 Affidabilità interoperatore sufficiente

4.3. RISULTATI - RACHIDE LOMBARE

I test più investigati, tra quelli risultati di adeguata qualità metodologica, sono stati quelli relativi alla palpazione delle spinose, utilizzando variegati approcci o combinazioni di tecniche. La palpazione è stata infatti presa in esame da 11 studi utilizzando diversi punti di repere (cresta iliaca, SIPS, palpazione avvalendosi del movimento), in differenti posizioni (prono, seduto, decubito laterale).

La mobilità è stata investigata in 8 studi, 2 dei quali hanno preso in esame la provocazione del dolore (Hidalgo e Schneider [16, 23]).

L'affidabilità è stata verificata in un numero maggiore di studi di buona qualità metodologica (16 per l'affidabilità, 5 per la validità), in diversi casi nel corso del medesimo studio. L'affidabilità interoperatore è stata testata più frequentemente di quella intraoperatore (14 e 4 volte rispettivamente). Non è stato usato un gold standard univoco per studiare la validità; al contrario quasi ogni studio ha utilizzato un reference standard diverso (RX in PA o laterale o entrambe, RMN dinamica, fluoroscopia).

Anche la tipologia di esaminatori è varia. In 8 studi hanno partecipato fisioterapisti, in altri studi hanno preso parte all'esame studenti, chiropratici, medici specializzati in terapia manuale, radiologi, anestesisti, radiologi specializzati in ecografia MSK.

La stessa esperienza clinica degli operatori è diversificata, anche perché 4 studi hanno ricercato il confronto tra operatori di esperienza diversa. In 4 studi il periodo di formazione non è stato specificato.

Rispetto al numero di pazienti oggetto di studio, due lavori hanno presentato meno di 10 pazienti. 10 studi sono stati svolti su un gruppo di 12-39 pazienti, 6 studi hanno esaminato tra i 51 e i 72 pazienti.

Pochi studi (4) hanno esaminato gruppi eterogenei di pazienti. La maggior parte ha selezionato sottopopolazioni di individui sani (8 studi) o con LBP (5 studi).

La randomizzazione dei pazienti è stata applicata in 8 studi su 16 che analizzavano l'affidabilità.

4.3.1. Identificazione dei punti di repere

La maggior parte degli studi presi in esame ha come oggetto la **palpazione dei reperi**. I risultati non sono facilmente confrontabili a causa di diversi approcci o combinazione di approcci e dei differenti obiettivi (confronto tra due tecniche palpatorie diverse, affidabilità intra-interesaminatore, correlazione tra palpazione e fattori legati all'esaminatore o ai soggetti).

La palpazione dei processi spinosi lombari ha esitato in una validità povera (47%, con valori tra 24% e 69%) per un solo studio di buona qualità (Harlick [7]) utilizzando un unico approccio a scelta, fosse esso adottando come riferimento la SIPS, la linea di Tuffier o tecniche di mobilità passiva. I valori salivano diventando moderatamente accurati (69%. $\kappa = 0.81$; CI 95%=0.79-0.83) invece con la combinazione di più tecniche (Snider [18]).

La validità è stata utilizzata anche per considerare i punti di repere più accurati da usare per l'identificazione delle spinose lombari.

Prendendo in esame l'affidabilità, nessuno studio di sufficiente qualità metodologica ha misurato la concordanza interoperatore nell'identificazione del livello vertebrale con tecniche di mobilità passiva.

La palpazione di L5 presenta una buona affidabilità (44% in stazione eretta, 38% da prono, che arriva a 89% e 88% per un intervallo inferiore a 10mm) sia inter (Cooper [1]) che intraoperatore ($k= 0.623$)(Merz [6]) utilizzando una combinazione di tecniche. Robinson [5] conferma la moderata affidabilità interoperatore(35% per repere preciso, 71% consentendo intervallo inferiore a 10mm, 87% per intervallo inferiore a 20mm) giustificandola col fatto che i terapisti dello studio lavoravano assieme ed utilizzavano procedure standardizzate. Anche Snider [18] rileva la medesima affidabilità interoperatore con 62% di accordo e $\kappa = 0.53$ (95% CI, 0.49-0.56), che diventa buona, col 73% di accordo e $\kappa = 0.66$ (95% CI, 0.61-0.71), se il confronto viene ristretto alla valutazione dei clinici rispetto al confronto con lo studente che ha partecipato allo studio .

Troke [15] attesta a differenza dei precedenti studi un'affidabilità intraoperatore quasi perfetta (ICC sul piano sagittale, coronale e orizzontale rispettivamente di 0.96, 0.80, e 0.98) e un'affidabilità interoperatore sostanziale (con l'86.4% di palpazioni accurate), spiegando questo dato al di sopra della media con l'omogeneità dei soggetti esaminati nello studio.

Un altro studio (Kim [9]) valuta l'affidabilità interoperatore in base al repere utilizzato per identificare i livelli vertebrali, e confrontando il margine inferiore della SIPS e la linea di Tuffier evidenzia una concordanza migliore nel gruppo che usava come repere la SIPS (K=0.92/0.96 per SISP, K=0.60/0.59 per cresta iliaca, P<0.05).

Ernst [22] mostra un alto accordo interoperatore (ICC>0.9) nell'identificazione delle spinose lombari utilizzando come strumento un curvilinee flessibile.

Esaminando i fattori correlati a validità ed affidabilità, Chakraverty [13] non rileva differenze tra operatori esperti ed inesperti (accuratezza rispettivamente del 57 e 54%), mentre Harlick [7] riporta differenze di accuratezza statisticamente significative tra operatori (OR=1.56, CI=1.10–2.23, P=0.01 per L3 e OR=1.74, CI=1.21–2.51, P<0.01 per L5), a prescindere dagli anni di esperienza.

I risultati dello studio Snider [18] sostengono che l'accuratezza dipende da esperienza (accuratezza del clinico 67%-78% vs 51% dello studente, P≤.03), anomalie anatomiche (anomalie della 12° costa P≤0.05), caratteristiche antropometriche del paziente (obesità P=.0003) .

4.3.2. Valutazione della mobilità

Gli studi che esaminano la **mobilità** riportano risultati meno positivi. La validità, testata in 3 studi, evidenzia povero accordo con Risonanza Magnetica (Landel [14]) nell'unico studio di buona qualità (Kappa=04,95% , CI=da -.16 a .24 per il segmento meno mobile, Kappa=00,95% CI=da-0.09 a 0.08, per il segmento più mobile).

L'affidabilità interoperatore risulta bassa/povera in 4 dei 6 studi che la testano (Bracht [24]: kappa da 0.12 a 0.30, Deore [19]: kappa tra 0.00 per L4 e 0.20 per L2, Johansson [12]: kappa pesato <0.20, Shneider [16]: kappa da −0.17 a 0.17) utilizzando pressione postero-anteriore per valutare il movimento intersegmentale o la rotazione. Landel [14] studia l'affidabilità interoperatore sul segmento più mobile e su quello ritenuto meno mobile, individuando rispettivamente povero (accordo 79.3%, Kappa=29.95%, CI=da -0.13 a 0.71) e buon accordo (accordo=82.8%, kappa=71,95%, CI=da 0.48 a 0.94).

L'affidabilità intraoperatore sembra invece eccellente da quanto rilevato da Shaw[21] (0.997, P<0.001), mentre lo studio di Deore [19] conferma miglior affidabilità intraoperatore che

interoperatore ma in ogni caso di livello povero (kappa intraoperatore tra 0.24 per L2 e 0.57 per L1).

L'affidabilità interoperatore aumenta nel momento in cui il parametro valutato sia il **dolore**. La provocazione del dolore dà risultati da discreti a buoni sia utilizzando come test i movimenti accessori, sia nei movimenti per identificare il pattern doloroso (kappa da 0.21 a 0.73) (Schneider [16]).

4.4. TABELLA SINOTTICA RACHIDE TORACICO

<table>
<tr><td colspan="6">1. Interrater reliability of a passive physiological intervertebral motion test In the mid-thoracic spine (Jean-Michel Brisme´e, 2006)</td></tr>
<tr><td>Distretto</td><td colspan="2">Obiettivo</td><td>Esaminatori</td><td>Soggetti</td><td>Risultati</td></tr>
<tr><td>Rachide toracico medio (T6-T7 o T5-T6) in sogg asintomatici</td><td colspan="2">Esaminare l'affidabilità inter esaminatore di un test di mobilità fisiologica intervertebrale passiva di un segmento di movimento medio-toracico della colonna vertebrale.</td><td>Tre fisioterapisti OMT con minimo 12 anni di esperienza clinica in terapia manuale.</td><td>41 studenti universitari (19 maschi e 22 femmine, tra i 19 e i 40 anni) sani, no interventi chirurgici, patologie reumatiche e gravidanza in anamnesi, BMI < 30</td><td>La percentuale di accordo per il test variava dal 63.4% al 82.5%. I punteggi di k variavano da 0.27 a 0.65 per un punteggio complessivo di 0.41.
Il test ha mostrato moderata l’affidabilità come definito da Landis e Koch.
L’indice di correlazione di Spearman (-0.44) non era statisticamente rilevante (P=21).
Il test PPIM quindi dimostra moderata affidabilità, quando eseguito su soggetti asintomatici</td></tr>
<tr><td colspan="6">AFFIDABILITA’ INTEROPERATORE: 5 (83%)</td></tr>
<tr><td>Order of observer randomized: 1</td><td>Case mix: 0</td><td>Observer blind to other observers findings: 1</td><td>Observer blind to counfounding informations: 1</td><td>Subjects blind to observer findings: 1</td><td>K/ICC: 1</td></tr>
</table>

2. Spinous process palpation using the scapular tip as a landmark vs a radiographic criterion standard. (Cooperstein, 2007)						
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Gold standard (validità)**	**Risultati**	
Rachide toracico	Determinare i punti di riferimento della colonna vertebrale in stazione eretta che corrispondono all' estremità inferiore della scapola e determinare l'accuratezza nel localizzare il processo spinoso (SP) 3 livelli sopra e sotto dal SP dato.	2 esperti chiropratici	34 soggetti , studenti di chiropratica, di cui il 41% femmine con età media 26 anni, asintomatici o minimamente sintomatici.	Rx standard	Deviazione standard 68% dei soggetti presentano l'estremità inferiore della scapola a livello di T7 T8 o T9. Un terapista manuale esperto può molto accuratamente localizzate 3 livelli vertebrali sopra e sotto al riferimento dato.	
AFFIDABILITA' INTEROPERATORE: Non sufficiente						
Order of observer randomized:	*Case mix:*		*Observer blind to other observers findings:*	*Observer blind to counfounding informations:*	*Subjects blind to observer findings:*	*K/ICC:*
VALIDITA' Quality score: 8						
Test validated against gold standard: 3	*Examiner blind to all applicable aspects: 0.5*	*Subject naive/without vested interest/unable affected result: 0.5*	*Description of test procedure and validation method: 1*	*Study subjects and examiners described: 1*	*Descriptions of statistics and and significance level defined: 1*	*Clear descriptio n protocol: 1*	*No additional problems: 0*

3. Reliability of joint mobility and pain assessment of the thoracic spine and rib cage in asymptomatic individuals (Heiderscheit, 2008)				
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Risultati**
Rachide dorsale e coste	determinare l'affidabilità inter- e intra-esaminatore della mobilità articolare della colonna vertebrale dorsale e gabbia toracica e della valutazione del dolore tra due esperti terapisti manuali.	Due terapisti manuali con oltre 10 anni di pratica clinica nelle tecniche di valutazione utilizzate.	9 soggetti sani (3 maschi e 6 femmine) senza storia di dolore dorso-lombare.	Mobilità articolare colonna dorsale: affidabilità intra-operatore compresa tra leggera e discreta (k= 0.17-0.26) basata sul calcolo dell'accordo ristretto. Affidabilità inter-operatore leggera (k=0.15). Secondo la definizione dell'accordo ampio l'affidabilità intra-esaminatore è aumentata per entrambi a livello buono (k=0.75/0.61) mentre l'affidabilità inter-esaminatore è aumentata al livello moderato (k=0.59). Dati simili si presentano per la gabbia toracica in quanto l'affidabilità intra-esaminatore aumenta da discreta a buona sotto l'accordo ampliato. L'affidabilità inter-esaminatore era consistentemente minore di quella intra-esaminatore (accordo ristretto k=0.11; accordo ampliato k=0.64). Provocazione del dolore: l'affidabilità intra- della colonna dorsale va da discreta a buona nell'accordo ristretto mentre sale a molto buono in quello ampio; l'affidabilità inter- era discreta ma sale a buona sempre nella versione ampliata. Per quanto riguarda la cassa toracica i valori di k sono da zero a buono nell'accordo ampio.
AFFIDABILITA' INTEROPERATORE: 5 (83%) INTRAOPERATORE: 3 (75%)				
Order of observer randomized: 1	*Case mix: 1*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*K/ICC: 1 accordo ampio*

4. Interexaminer reliability of thoracic motion palpation using confidence ratings and continuous analysis (Cooperstein, 2010)				
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Risultati**
T7 – T10	Determinare se consentire all'esaminatore di valutare la propria fiducia nelle sue conclusioni, anche utilizzando un metodo analitico di dati continui, avrebbe influenzato il livello di concordanza.	Due esaminatori con titolo di chiropratico, ciascuno con più di 20 anni di esperienza clinica	I soggetti erano 52 studenti di chiropratica, asintomatici. I soggetti sono stati esclusi se presente a metà schiena dolore più grande di 2 sulla scala del dolore numerica a 11 punti o se non potevano tollerare la procedura di palpazione per qualsiasi altro motivo.	Divisione in sottogruppi. Per tutti i soggetti, le risposte degli esaminatori erano "scarso" (gruppo 1): ICC [2,1] = 0,3110 (95% CI, 0,0458-0,5358). Al contrario, un accordo inter-esaminatore era "buono" quando entrambi gli esaminatori erano molto fiduciosi (gruppo 2): ICC [2,1] = 0,8266 (95% CI 0,6257-0,9253). Quando ogni esaminatore era "molto fiducioso" sul segmento toracico fissato, i livelli che sono stati identificati erano molto vicini (media di differenza 2 cm, circa un livello vertebrale; nel gruppo 6 con ICC=−0.3873 (95% CI, −0.7907 - 0.2729) la differenza media era 7.10 cm). Ciò corrisponde a un accordo "buono", un risultato non comune nella maggior parte degli studi di palpazione del movimento inter-esaminatore. Pertanto, il livello di fiducia degli esaminatori ha avuto un effetto sull'affidabilità inter-esaminatore della colonna vertebrale toracica.
AFFIDABILITA' INTEROPERATORE: 4 (66%)				
Order of observer randomized: 1	*Case mix: 0*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 0*	*K/ICC: 1*

5. Determination of thoracic and lumbar spinal processes by their percentage position between C7 and the PSIS level (Ernst, 2013)				
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Risultati**
T4,T7,T10 L1, L4	Verificare l'affidabilità inter-operatore nel rilevare la posizione dei processi spinosi di T4, T7, T10, L1, L4 come valore percentuale della distanza tra C7 e la SIPS, utilizzando un righello flessibile.	6 fisioterapisti che effettuavano la misurazione a coppie (queste erano formate a seconda delle disponibilità degli esaminatori)	22 soggetti sani tra 22 e 51 anni senza evidenti deviazioni morfologiche del rachide.	Affidabilità relativa quasi perfetta: ICC > 0.9. Affidabilità assoluta presenta un basso errore standard: SEM = 0.2/0.3
AFFIDABILITA' INTEROPERATORE: 3 (50%)				
Order of observer randomized: 0	*Case mix: 0*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 0*	*K/ICC: 1*

6. Mapping intended spinal site of care from the upright to prone position: an interexaminer reliability study. (Cooperstein, 2014)							
Distretto	**Obiettivo**		**Esaminatori**	**Soggetti**	**Gold standard (validità)**		**Risultati**
Rachide toracico (focus on T7)	Indagare come le associazioni di punti di riferimento spino-scapolari possono condurre a errori nel trattamento e nella mappatura dei livelli spinali.		Due esperti con anni di esperienza d'insegnamento e pratica clinica e 8 studenti novizi.	2 soggetti uomini sani	Rx da prono		ICC Errore di 3 livelli dall'individuazione in stazione eretta alla posizione prona, in particolare la palpazione del movimento tradurrebbe in un errore di mira di 2 livelli per l'intervento e di 3 per la creazione di grafici. L'affidabilità dimostra che è stato mantenuto accuratamente il contatto con lo stesso processo spinoso mentre il soggetto andava dalla posizione eretta a quella prona. ICC(2,1)= 0.83
AFFIDABILITA' INTEROPERATORE: 5 (83%)							
Order of observer randomized: 1	*Case mix: 1*		*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*Subjects blind to observer findings: 0 non specificato*		*K/ICC: 1*
VALIDITA' Quality score: 9.5							
Test validated against gold standard: 3	*Examiner blind to all applicable aspects: 1*	*Subject naive/without vested interest/unable affected result: 1*	*Description of test procedure and validation method: 1*	*Study subjects and examiners described: 1*	*Descriptions of statistics and significance level defined: 0.5*	*Clear description protocol: 1*	*No additional problems: 1*

7. Interrater Reliability of Motion Palpation in the Thoracic Spine (Bruce F. Walker, 2015)				
Distretto	**Obiettivo**	**Esaminatori**	**Soggetti**	**Risultati**
Rachide toracico in soggetti eterogenei	Indagare l'affidabilità inter-operatore della palpazione del movimento della colonna dorsale per percepire restrizioni articolari e dolore.	Due chiropratici esperti cha hanno utilizzato sia il loro approccio pragmatico che la tecnica standardizzata. Entrambi hanno lavorato in cieco, sia tra essi che nei confronti dei soggetti.	25 soggetti di età compresa tra 18 e 70 anni, di cui 5 con dolore dorso-lombare il giorno della valutazione e 18 hanno avuto dolore dorso-lombare in precedenza.	I valori di k per la restrizione segmentale variavano tra -0.27 e 0.36 a seconda del livello spinale. Anche per quanto riguarda la provocazione del dolore i valori di k sono deludenti: k era compreso tra -0.38 e 0.32, a seconda del livello spinale. Il punteggio della tecnica pragmatica era peggiore rispetto a quella standardizzata.
AFFIDABILITA' INTEROPERATORE: 6 (100%)				
Order of observer randomized: 1	*Case mix:1*	*Observer blind to other observers findings: 1*	*Observer blind to counfounding informations: 1*	*K/ICC: 1*

4.5. TABELLA DEGLI STUDI ESCLUSI PER INSUFFICIENTE QUALITA' METODOLOGICA – RACHIDE TORACICO

Titolo	**Autore, anno, rivista**	**Target (argomento e distretto)**	**Punteggio qualità**
Spinous process palpation using the scapular tip as a landmark vs a radiographic criterion standard.	Cooperstein, 2007 *Journal of Chiropractic Medicine*	Validità RT (riferimenti di individuazione dei livelli spinali) Affidabilità RT (localizzazione delle spinose)	2/6 * *Non sufficiente solo per affidabilità
Simple anatomical information improves the accuracy of locating specific spinous processes during manual examination of the low back.	Philips, 2009 *Manual Therapy*	Validità RL (L3) e RT (T12)	6/10

4.6. RISULTATI RACHIDE TORACICO

In seguito alla valutazione che ha permesso di prendere in considerazione solo studi di elevata qualità metodologica presentati negli ultimi 10 anni, è emerso che nell'ambito del rachide toracico sono stati analizzati complessivamente in egual numero sia le tecniche di mobilità passiva che quelle di palpazione delle spinose. La mobilità intervertebrale è stata infatti oggetto di studio di 3 lavori [30, 31, 32] di cui due comprendevano anche la provocazione del dolore (Heiderscheit e Walker) mentre 4 studi si sono interessati dell'esame palpatorio prendendo in considerazione differenti aspetti che concernono la rilevazione delle spinose, quali l'utilizzo di reperi anatomici (C7, angolo inferiore della scapola e SIPS) e i relativi errori nella localizzazione dei livelli vertebrali [33, 34, 34] o la possibile influenza che può avere il livello di fiducia dell'esaminatore nella tecnica effettuata [36].

Prendendo in considerazione solo i valori che hanno raggiunto i requisiti di punteggio secondo la checklist e il cutoff di Haneline [11] per l'affidabilità (Fig. 2] e quelli di Hestbaek [10] per la validità (Fig. 1), la prima risulta maggiormente studiata. Troviamo infatti il target validità solo in 2 studi, i quali hanno utilizzato il medesimo gold standard (RX da prono o in posizione non altrimenti specificata), a fronte dei 5 studi relativi all'affidabilità, di cui solamente uno trattava quella intra-operatore oltre a quella inter-esaminatore. Solamente uno studio non ha superato i cut off prestabiliti [25], in particolare quelli relativi all'affidabilità non sono stati raggiunti nemmeno dallo studio di [33], il quale però è risultato rilevante in termini di validità.

Gli esaminatori vedono una varietà di partecipazione in termini di figure professionali e relative specializzazioni: 3 studi hanno utilizzato fisioterapisti di cui 2 con titolo di terapista manuale mentre gli altri 4 studi hanno visto la partecipazione di chiropratici. Il numero degli esaminatori scelti era principalmente 2, ma questo poteva variare salendo a 3 [30] o 6 [22] componenti. La pratica clinica, specificata nella maggioranza dei casi, contava almeno 10 anni di esperienza, fatta eccezione per l'esame di [36] in cui due professionisti con rilevante expertise si sono interfacciati a 8 studenti novizi.

I soggetti reclutati in questi studi non si presentavano completamente eterogenei in quanto età, sesso, altezza e peso erano posti in un range abbastanza ampio ma tutti risultavano sani o asintomatici, tranne che per lo studio di [32] in cui 23 individui su 25 attestavano la presenza di dolore dorso-lombare al momento dell'esame o in precedenza. Il numero dei soggetti era

compreso tra 22 e 41, tranne che per due studi che presentavano un campione notevolmente più ristretto di 2 [34] e 9 [31] componenti.

L'analisi del metodo con cui sono stati condotti gli studi inclusi ha evidenziato in alcuni casi delle carenze nella descrizione della procedura, infatti non è stato possibile soddisfare alcuni items delle checklist utilizzate in quanto il criterio non veniva specificato nel testo. Ciò si riferisce in particolare alla modalità in cieco da parte degli esaminatori o dei soggetti relativamente ai risultati delle tecniche o a eventuali fattori di contesto.

Tuttavia i punteggi calcolati hanno permesso di assegnare agli studi un' elevata qualità: l'affidabilità inter-operatore si colloca complessivamente nella fascia medio-alta di 83%, che vede un picco massimo di ottima qualità con lo studio di [32] che soddisfa tutti gli items mentre il valore minimo è di [22] che ottiene comunque una buona qualità (50%) e presenta inoltre un'affidabilità relativa pressoché perfetta (ICC>0.9) così come quella assoluta data da un alto accordo intra-operatore. Anche l'unico lavoro preso in considerazione sull'affidabilità intra-operatore [31] presenta un buon livello (75%). Il medesimo discorso vale per la validità, molto buona in quanto i lavori di Cooperstein registrano punteggi di 8/10 [33] e 9.5/10 [34].

Considerando i risultati dei singoli studi in relazione alle tecniche indagate, risulta complicato confrontarne i dati in quanto essi, pur nello stesso ambito e con lo stesso quesito, si presuppongono obiettivi con connotazioni differenti e presentano approcci con procedure e riferimenti anch'essi diversi.

4.6.1 Identificazione dei punti di repere

Per quanto riguarda l' **esame palpatorio** uno degli intenti più ricercati è stato determinare i livelli della colonna vertebrale che corrispondono ai vari reperi anatomici; in particolare Cooperstein, in un suo studio [33] di buona validità (8/10), si è focalizzato sull'estremità inferiore della scapola che nel 68% dei soggetti si trova a livello T7, T8 o T9 e ha tentato di determinare l'accuratezza con cui un terapista manuale esperto è in grado di localizzare 3 livelli vertebrali sopra e sotto al riferimento dato. Tuttavia non è stato possibile tenere in considerazione quest'ultima affermazione a causa del mancato raggiungimento del cut off di affidabilità.

Lo stesso autore si è posto il dubbio di indagare ulteriori differenti aspetti della tecnica di palpazione ovvero come le associazioni dei reperi spino-scapolari possano condurre a errori sia nel trattamento che nella mappatura dei livelli spinali. Il lavoro, di elevata validità (8.5/10) ha confermato un errore di 3 livelli dall'individuazione in stazione eretta alla posizione prona, in particolare la palpazione del movimento tradurrebbe in un errore di mira di 2 livelli per il trattamento e di 3 per la creazione di grafici. Con il secondario obiettivo di verificare un nuovo metodo per la mappatura dei siti spinali durante il passaggio dalla stazione eretta alla posizione prona, è stata esaminata l'affidabilità inter-esaminatore, la quale dimostra che è stato mantenuto accuratamente il contatto con lo stesso processo spinoso mentre il soggetto si portava dalla posizione eretta a quella prona (ICC = 0.83). Anche [22] si è occupato della rilevazione dei livelli vertebrali, in particolare dei processi spinosi di T4, T7, T10 riportati come valore percentuale della distanza tra C7 e SIPS, raggiungendo un'affidabilità relativa quasi perfetta (ICC>0.9) ed un alto accordo intra-operatore. Viene assegnato il valore percentuale di 39% e 54,1% rispettivamente a T7 e T10, livello quest'ultimo che rappresenta il valore massimo con 2.6% tra gli estremi delle misurazioni.

Ciò che può ulteriormente influire sull'affidabilità inter-esaminatore della palpazione della colonna toracica è il livello di fiducia degli esaminatori sulla propria fissazione del livello vertebrale. Questo viene affermato da [36], il cui studio, oltre ad essere di buona affidabilità, presenta un accordo inter-esaminatore buono quando entrambi gli esaminatori si giudicavano "molto fiduciosi" (ICC = 0.83), il che è un risultato non comune nella maggior parte degli studi sulla palpazione del movimento inter-esaminatore; in questa condizione i livelli identificati risultavano molto vicini (massima differenza di 2 cm).

4.6.2 Valutazione della mobilità

La **mobilità intervertebrale** è stata studiata da 3 autori: [30, 31, 32], i quali si sono interessati ciascuno di aspetti che differivano in alcuni particolari.

La mobilità fisiologica intervertebrale passiva di un segmento medio-toracico, studiata da Brisme'e [30] presenta, a fronte di un indice di correlazione non statisticamente rilevante (p=21), una moderata affidabilità inter-esaminatore (k= 0.41) quando la tecnica viene eseguita solamente su soggetti asintomatici. Considerando la mobilità in termini di restrizione articolare della colonna dorsale, lo stesso livello viene raggiunto da Hederscheit [31],

applicando però un'estensione dell'accordo che porta appunto l'affidabilità inter-esaminatore da lieve a moderata e quella intra-esaminatore da discreta a buona. Lo stesso si verifica prendendo in esame la gabbia toracica, in cui l'affidabilità intra-esaminatore cresce da discreta a buona applicando l'accordo ampio. Il medesimo studio ha incluso una **valutazione del dolore** che a livello del rachide toracico mostra un'affidabilità intra-esaminatore discreta/buona mentre nella gabbia toracica i valori di k sono pari a 0, tuttavia questi diventano rispettivamente molto buoni e buoni sempre nella versione ampliata. Complessivamente nel lavoro di Hederscheit [31] l'affidabilità inter-esaminatore era consistentemente minore di quella intra-esaminatore (accordo ristretto k=0.11; accordo ampliato k=0.64). Ovviamente i dati risultano differenti a seconda del livello spinale testato, infatti nello studio di Walker [32], la cui qualità è elevata avendo soddisfatto tutti gli items della checklist per l'affidabilità inter-esaminatore, i valori di k per la restrizione segmentale variavano tra -0.27 e 0.36 a seconda del livello spinale. Coerentemente ai risultati dell'autore precedente, nell'ambito dell'accordo ristretto, la **provocazione del dolore** mostra valori di k deludenti: k era compreso tra -0.38 e 0.32, sempre a seconda del livello spinale. Un'altra variabile che complica il confronto tra i lavori è, come già accennato, l'utilizzo di tecniche differenti: in quest'ultimo studio infatti, a differenza degli altri, è stata comparata una tecnica standardizzata ad una pragmatica (di utilizzo propria del singolo esaminatore), la quale ha ottenuto un punteggio peggiore. Non sono stati evidenziati studi che abbiano svolto indagini sulla validità delle tecniche di mobilità intervertebrale nel rachide toracico.

In seguito alla valutazione che ha permesso di prendere in considerazione solo studi di elevata qualità metodologica presentati negli ultimi 10 anni, è emerso che nell'ambito del rachide toracico sono stati analizzati complessivamente in egual numero sia le tecniche di mobilità passiva che quelle di palpazione delle spinose. La mobilità intervertebrale è stata infatti oggetto di studio di 3 lavori (Brisme'e, Heiderscheit e Walker) di cui due comprendevano anche la provocazione del dolore (Heiderscheit e Walker) mentre 4 studi si sono interessati dell'esame palpatorio prendendo in considerazione differenti aspetti che concernono la rilevazione delle spinose, quali l'utilizzo di reperi anatomici (C7, angolo inferiore della scapola e SIPS) e i relativi errori nella localizzazione dei livelli vertebrali (Cooperstein, 2007 e 2014 ed Ernst, 2013) o la possibile influenza che può avere il livello di fiducia dell'esaminatore nella tecnica effettuata (Cooperstein, 2010).

Prendendo in considerazione solo i valori che hanno raggiunto i requisiti di punteggio secondo la checklist e il cutoff di Haneline (2009) per l'affidabilità e quelli di Hestbaek (2000) per la

validità, la prima risulta maggiormente studiata. Troviamo infatti il target validità solo in 2 studi, i quali hanno utilizzato il medesimo gold standard (RX da prono o in posizione non altrimenti specificata), a fronte dei 5 studi relativi all'affidabilità, di cui solamente uno trattava quella intra-operatore oltre a quella inter-esaminatore.

Gli esaminatori vedono una varietà di partecipazione in termini di figure professionali e relative specializzazioni: 3 studi hanno utilizzato fisioterapisti di cui 2 con titolo di terapista manuale mentre gli altri 4 studi hanno visto la partecipazione di chiropratici. Il numero degli esaminatori scelti era principalmente 2, ma questo poteva variare salendo a 3 (Brisme'e, 2006) o 6 (Ernst, 2013) componenti. La pratica clinica, specificata nella maggioranza dei casi, contava almeno 10 anni di esperienza, fatta eccezione per l'esame di Cooperstein (2010) in cui due professionisti con rilevante expertise si sono interfacciati a 8 studenti novizi.

I soggetti reclutati in questi studi non si presentavano completamente eterogenei in quanto età, sesso, altezza e peso erano posti in un range abbastanza ampio ma tutti risultavano sani o asintomatici, tranne che per lo studio di Walker (2015) in cui 23 individui su 25 attestavano la presenza di dolore dorso-lombare al momento dell'esame o in precedenza. Il numero dei soggetti era compreso tra 22 e 41, tranne che per due studi che presentavano un campione notevolmente più ristretto di 2 (Cooperstein, 2014) e 9 (Heiderscheit, 2008) componenti.

L'analisi del metodo con cui sono stati condotti gli studi inclusi ha evidenziato in alcuni casi delle carenze nella descrizione della procedura, infatti non è stato possibile soddisfare alcuni items delle checklist utilizzate in quanto il criterio non veniva specificato nel testo. Ciò si riferisce in particolare alla modalità in cieco da parte degli esaminatori o dei soggetti relativamente ai risultati delle tecniche o a eventuali fattori di contesto.

Tuttavia i punteggi calcolati hanno permesso di assegnare agli studi un'elevata qualità: l'affidabilità inter-operatore si colloca complessivamente nella fascia medio-alta di 83%, che vede un picco massimo di ottima qualità con lo studio di Walker (2015) che soddisfa tutti gli items mentre il valore minimo è di Ernst (2013) che ottiene comunque una buona qualità (50%) e presenta inoltre un'affidabilità relativa pressoché perfetta (ICC>0.9) così come quella assoluta data da un alto accordo intra-operatore. Anche l'unico lavoro preso in considerazione sull'affidabilità intra-operatore (Heiderscheit, 2008) presenta un buon livello (75%). Il medesimo discorso vale per la validità, molto buona in quanto i lavori di Cooperstein registrano punteggi di 8/10 (2007) e 9.5/10 (2014).

5. DISCUSSIONE

5.1. RACHIDE LOMBARE

La necessità di indagare gli aspetti correlati all'esame palpatorio e di mobilità passiva per la gestione delle disfunzioni legate al rachide lombare rimane da anni un argomento di discussione che interessa la terapia manuale. Tali dati infatti vengono ritenuti punto di partenza per il processo decisionale che guida la scelta degli interventi terapeutici e vengono relazionati alle misure di disabilità del paziente in carico.
Sebbene sia presente un buon numero di pubblicazioni sull'argomento, i risultati permangono controversi e le inferenze in pratica clinica, come ad esempio l'abitudine di correlare diminuzione di ROM con presenza di disabilità, rimangono prive di solidi elementi sulla validità [26].
Ciò nonostante, tali test vengono ad oggi utilizzati frequentemente nella pratica clinica dei terapisti manuali, pur essendo il loro impiego supportato principalmente dall'opinione di esperti, parametro però attualmente insufficiente per dare validità di contenuto per la valutazione e gestione delle disfunzioni vertebrali.

La stessa percezione di accuratezza palpatoria da parte dei terapisti, in linea con la scarsa disponibilità di misure oggettive di validità in relazione alla frequenza d'utilizzo dei suddetti approcci manuali, sembra evidenziare una non piena sicurezza degli operatori nella valutazione delle restrizioni del rachide, con una diminuzione della stessa nell'approccio al rachide lombare rispetto alle altre regioni vertebrali; sicurezza che si incrementa con l'esperienza così come aumenta l'utilizzo di tecniche di trattamento che prevedono una precisione manuale propedeutica, quali le manipolazioni [27].

L'esame di mobilità è stato indagato in altri distretti per verificare il miglioramento della restrizione nel rachide cervicale post manipolazione, a riprova dell'importanza, nella cultura del terapista manuale, dell'end feel come parametro di classificazione delle disfunzioni e valutazione dell'efficacia di trattamento [28].

Per questi motivi è importante continuare a fare chiarezza sulla consistenza di tali procedure, in primis concretizzandone la validità e l'affidabilità, così da procedere a dare loro il giusto peso nella gestione del ragionamento in pratica clinica.

I riscontri della valutazione metodologica alla base della selezione degli articoli per questa revisione sistematica mostrano che affidabilità e validità e sono state verificate in egual misura (16 per l'affidabilità, 15 per la validità), in diversi casi nel corso del medesimo studio.
Solo 5 studi di validità hanno però superato il cut off rappresentato da un punteggio ≥80%, per molteplici motivi, tra cui l'eterogeneità dei reference standard presi a riferimento ed alcune mancate descrizioni che non hanno permesso di concedere punti.
Riguardo all'affidabilità, 6 studi analizzanti l'affidabilità interoperatore hanno registrato buona qualità (>50%), 8 studi hanno raggiunto il 50% del punteggio complessivo, mentre due studi sono stati definiti di scarsa qualità. Tutti e 5 gli studi che verificavano l'affidabilità intraoperatore sono risultati di buona qualità.
La qualità metodologica degli studi non è correlata ai test esaminati (palpazione, esame della mobilità o provocazione del dolore) né alla tipologia di studio (affidabilità, validità).
La validità è stata verificata solo in tre studi sulla mobilità/provocazione del dolore, uno di buona qualità e due di livello non sufficiente. I restanti studi sulla validità avevano come oggetto l'esame palpatorio e sono per la metà di buon livello (7 studi), per l'altra di scarsa qualità metodologica (6 studi).
L'affidabilità interoperatore è risultata da sufficiente a buona in tutti e 5 gli studi sulla mobilità/provocazione del dolore, insufficiente in due degli 11 studi che esaminavano test palpatori. L'affidabilità intraoperatore è stata definita di buon livello sia nei 2 studi che esaminavano la mobilità sia nei due sulla palpazione.

Alcuni articoli hanno citato possibili bias all'interno del protocollo di studio. L'ordine non randomizzato di pazienti (in 8 studi su 16 che analizzavano l'affidabilità) non elimina possibili bias. L'uso di strumentazione per la valutazione della validità in un caso prevedeva l'utilizzo di un unico ecografo con un solo operatore, comportando un possibile bias di misura (Shaw [21]), mentre in un altro studio per l'affidabilità interoperatore non prevedeva cecità ai risultati della prima fase, con un possibile bias di risultati (Chakraverty [13]). Questi tra i motivi che non hanno permesso ai suddetti lavori di superare il cut off rispettivamente di validità ed affidabilità interoperatore e di essere inclusi solo parzialmente in quanto idonei relativamente ad affidabilità interoperatore e validità. Anche l'utilizzo di diversi test per sessione ed in ordine non specificato può essere fonte di bias di risultati.

La descrizione di procedura d'esame e soggetti non sempre è stata soddisfacente. Non tutti i risultati sono stati esposti in modo chiaro e preciso. Talvolta la stessa presentazione, confusa o parzialmente lacunosa, ha reso difficile il confronto tra studi, oltre all'utilizzo di modalità d'analisi statistiche spesso ma non sempre uguali per tutti (coefficiente k o K pesato o ICC per affidabilità, o semplicemente valori percentuali).
In tutti gli studi inerenti la validità escluso il lavoro di Landel [14] è stato presentato il livello di significatività statistica (p value), mentre in quasi tutti gli studi di affidabilità sono stati identificati indici d'accordo quali K, K pesato, ICC.
Le conclusioni riportate nei diversi studi sono supportate dai dati presenti nelle tabelle e rispecchiano i requisiti metodologici degli stessi, già esplicitati nei paragrafi precedenti.

Per quanto concerne i risultati emersi, come enunciato dalle precedenti revisioni sistematiche (Hestbaek [10], Stockendhal [29]), si conferma la maggior affidabilità degli studi che testano la provocazione del dolore, mentre la valutazione della mobilità presenta minore affidabilità, comunque non molto al di sotto di quella riassunta dagli studi aventi per oggetto l'esame palpatorio.
A prescindere dall'obiettivo dello studio, l'affidabilità interoperatore si è dimostrata inferiore al test-retest.
Non si possono trarre considerazioni chiare dagli studi di validità raffrontando test palpatori e di mobilità, visto il numero esiguo di lavori di buona qualità soprattutto sull'esame della mobilità. Come già sottolineato, viene segnalata la mancanza di un gold standard univoco che avrebbe permesso un confronto almeno tra gli studi sull'esame palpatorio. Quasi ogni studio ha utilizzato un reference standard diverso: oltre ai citati esami strumentali degli studi di buona qualità quali RX in PA o laterale, fluoroscopia, RMN dinamica, studi di scarso livello metodologico hanno impiegato ultrasuoni, RMN posizionale, fino ad arrivare alla palpazione di esperti o alla semplice analisi di Sensibilità-Specificità.
Inoltre è necessario considerare che il livello di accuratezza dipende dall'intervallo considerato, e non esiste una misura comune negli studi [5]. Come detto è confermata la maggior consistenza dell'affidabilità per gli studi che considerano la provocazione del dolore, ma permangono problemi a dimostrarne la validità: non vi è possibilità di comprovare infatti che il dolore sia correlato a disfunzione. Nonostante ciò rimane un parametro guida quanto mai necessario per localizzare l'area di trattamento e dosare le tecniche manuali selezionate in base all'obiettivo definito a seguito della valutazione anamnestica e dell'esame fisico.

I risultati inoltre non sono facilmente confrontabili perché adottano diverse tecniche e posizioni, anche tra esse associate. Nel caso in cui vengano usati test differenti durante la medesima sessione sullo stesso paziente, il primo influenza l'interpretazione dell'ultimo, quindi non si possono fare riflessioni sull'accuratezza dei singoli test quanto sul fatto che una combinazione di test permette di raggiungere un maggior valore di accuratezza.
Si vuole inoltre segnalare che il cambiamento di posizione che intercorre tra l'esame palpatorio e l'imaging che ne testa la validità può essere fonte d'imprecisione a causa del probabile spostamento della pelle. Alcuni studi hanno cercato di ovviare al problema con un controllo del posizionamento del marker prima dell'imaging o con un esame palpatorio direttamente nella posizione dell'esame strumentale. Bisogna considerare però che nella clinica alcune posizioni di valutazione sono diverse dalle procedure utilizzate negli studi, pertanto i dati ricavati dai suddetti lavori non possono essere trasferiti a posizioni differenti da quelle considerate nello studio iniziale.
Il limitato peso dei dati di validità ed affidabilità si associa ad una debole e confusa correlazione con i parametri legati all'esperienza dell'esaminatore od alle caratteristiche antropometriche del soggetto valutato. Tutto ciò avvalora l'ipotesi che il valore del test dipenda soprattutto da abilità intrinseche all'esaminatore stesso, e in quanto tali non soggette a categorizzazioni oggettivabili.

Ulteriori problemi nell'interpretazione dei risultati sono dovuti al fatto che molti studi hanno preso in esame soggetti asintomatici, o escluso nella selezione anomalie strutturali, definendo un sottogruppo di individui che si allontana dall'eterogeneità della popolazione reale.
Oltre al fatto che nella pratica clinica l'esaminatore deve tenere in considerazione di poter incontrare pazienti che presentano anomalie strutturali, l'esclusione dal gruppo valutato di tali presenze unito alla scelta di utilizzare soggetti sani porta al rischio di inficiare i dati di validità ed affidabilità non essendo il campione rappresentativo della popolazione.
Alcuni studi inoltre hanno riportato risultati in percentuale senza calcolare il grado di accordo e senza definire il livello di significatività statistica.

Calando le riflessioni in pratica clinica, nonostante i risultati non siano facilmente paragonabili per l'eterogeneità degli studi considerati, sembra che le caratteristiche di ogni singolo lavoro (competenze e livello di esperienza degli esaminatori, consenso e training nelle procedure usate, valutazione di soggetti sintomatici o sani) abbiano minima influenza sui risultati, che attestano poveri valori di affidabilità e validità di esame palpatorio e di mobilità.

I test in parallelo (l'associazione di più tecniche palpatorie e di mobilità) aumentano la riproducibilità e, come avviene nella consuetudine lavorativa, permettono all'esaminatore di prendere più facilmente una decisione procedurale per l'intervento terapeutico. Il problema maggiore è che diventa un processo molto complesso stilare una batteria di test in sequenza da applicare in pratica clinica per aumentarne l'affidabilità, mancando peraltro dati che ne sostengano la validità.

Rigidi parametri di validità interna tra l'altro rischiano di diminuire l'applicabilità delle procedure in pratica clinica e conseguentemente diminuirne la validità esterna.
Ad ogni modo escludere più della la metà degli studi esaminati per insufficiente qualità metodologica di validità dimostra la forte importanza di curare questo aspetto nella produzione di lavori futuri, pur tenendo presente la difficoltà di dimostrare l'accuratezza dei test esaminati che però, come sottolineato, prescinde dal buon livello metodologico che è necessario dare ai lavori pubblicati.
E' auspicabile pertanto che gli studi vengano direzionati verso una selezione più eterogenea di soggetti, che gli intervalli di accuratezza siano definiti in modo preciso e standardizzato per ottenere dati confrontabili, che l'impatto dei criteri di inclusione venga correlato con i risultati ottenuti e sia specificata la qualità metodologica dei lavori portati avanti.

5.2. RACHIDE TORACICO

La colonna dorsale è considerata una zona di transizione tra le regioni spinali cervicali e lombari. L'associazione tra le vertebre della regione toracica e le costole corrispondenti ne identifica il ruolo di stabilità, distribuzione del carico e respirazione. Recenti acquisizioni considerano il rachide dorsale come una fonte di dolore e confermano l'influenza che può avere la sua mobilità sul resto della colonna e nel cingolo scapolare; nondimeno alcuni studi hanno confermato il successo del trattamento manuale della colonna toracica in numerose condizioni cliniche. Anche se l'evidenza che supporta l'uso della mobilizzazione e manipolazione nel rachide toracico continua a incrementare, quella che attesta l'abilità di individuare livelli specifici della colonna rimane controversa [37]. Nel processo di identificazione dei livelli spinali inoltre è importante tenere presente anche la variabilità anatomica tra i pazienti, da un lato, e un certo margine di errore da parte dell'esaminatore dall'altro [33].

L'affidabilità inter-operatore della valutazione della mobilità della colonna però solo raramente supera i valori di probabile correttezza e studi di affidabilità sono impegnativi per condotta e spesso difficili da interpretare a causa dei loro difetti metodologici e statistici [30]. La medesima situazione si verifica nell'ambito dei test di provocazione del dolore [31].

Sono frequenti infatti i lavori in cui si riscontrano carenze metodologiche, quali la non cecità degli esaminatori, ad esempio, e l'analisi del rapporto inter- o intra- esaminatore. In alcuni casi vengono trascurate anche le condizioni di esecuzione di queste tecniche, da prono o in stazione eretta ad esempio, abbassando il livello di accuratezza ed esitando in scarsa affidabilità. In particolare differenze nella tecnica di palpazione sembrano essere la fonte più probabile di bassa affidabilità inter-esaminatore. Altri autori hanno suggerito che le difficoltà nell'individuazione dei reperi anatomici e l'uso di soggetti asintomatici potrebbero anch'esse essere responsabili della riduzione di affidabilità [39].

Essendo oggetto di queste problematiche, le ricerche si sono concentrate sull'analisi dei criteri di qualità degli studi in tema, eseguendo valutazioni metodologiche orientate a verificarne gli specifici gradi di validità e affidabilità. Una revisione della letteratura sull'affidabilità della palpazione spinale statica conferma infatti un basso grado di accordo, sia inter- che intra-esaminatore, fra i tre metodi di palpazione statica comunemente utilizzati in terapia manuale

[11] sia da studenti che da clinici esperti tanto da suggerire che l'uso continuato di qualsiasi test diagnostico come metodo per il rilevamento di disfunzioni e scelta di interventi è di dubbia utilità [40].

Cosciente di queste evidenze e con il l'intento di fare maggiore chiarezza in questo ambito, tentando di affinare i criteri di ricerca e analisi delle precedenti revisioni per colmare eventuali lacune metodologiche, lo studio in questione ha analizzato nell'ambito del rachide toracico complessivamente 7 studi che hanno superato il cut off d'inclusione metodologica, riguardanti in egual numero sia le tecniche di mobilità passiva che quelle di palpazione delle spinose.

I punteggi calcolati hanno permesso di assegnare agli studi un'elevata qualità: l'affidabilità inter-operatore si colloca complessivamente nella fascia medio-alta di 83%, che vede un picco massimo di ottima qualità con lo studio di Walker [32] che soddisfa tutti gli items mentre il valore minimo è di Ernst [22] che ottiene comunque una buona qualità (50%) e presenta inoltre un'affidabilità relativa pressoché perfetta (ICC>0.9) così come quella assoluta data da un alto accordo intra-operatore. Anche l'unico lavoro preso in considerazione sull'affidabilità intra-operatore [31] presenta un buon livello (75%).

Il medesimo discorso si ha per la validità, molto buona in quanto i lavori di Cooperstein registrano punteggi di 8/10 [33] e 9.5/10 [34]. Solamente uno studio non ha superato i cut off prestabiliti [25], in particolare quelli relativi all'affidabilità non sono stati raggiunti nemmeno dallo studio di [33], il quale però è risultato rilevante in termini di validità.

Si è riscontrata una maggiore affidabilità negli studi che testano la mobilità, specie in associazione con la provocazione del dolore [32, 31], in particolare risulta più rappresentata l'affidabilità inter-esaminatore. Gli studi invece che riguardano l'affidabilità delle tecniche di fissazione risultano più mediocri.

I risultati di questi studi sono difficilmente comparabili in quanto trattano aspetti diversi della stessa tecnica.

Per quanto riguarda la validità troviamo un buon livello negli studi, tra l'altro dello stesso autore (Cooperstein) [33, 34, 36], che indagano la palpazione e la rilevazione dei livelli spinali con relativi reperi per permettere una possibile mappatura. In questo caso, piuttosto che confrontare gli studi, è opportuno tenerli in considerazione in maniera complementare in quanto ne emerge che uno dei reperi più utilizzati, l'angolo inferiore della scapola, risulta

collocarsi nel 68% dei soggetti a livello T7, T8 o T9 e, nonostante un terapista manuale esperto possa accuratamente localizzare 3 livelli sopra e 3 livelli sotto il riferimento dato, ciò che può inficiare la rilevazione sono la posizione del paziente e l'influenza del pollice dell'esaminatore in termini di posizionamento. Nel secondo studio infatti emerge che vi può essere un errore di almeno 3 livelli nell'individuazione dalla stazione eretta alla posizione prona, in particolare di 2 livelli in sede di trattamento e 3 per quanto concerne la rilevazione per la creazione di grafici e mappature. Tuttavia l'analisi dell'affidabilità associata a questa tecnica rivela che ha un buon valore per il mantenimento del contatto nel passaggio di posizione (ICC = 0.83).

Il problema nella considerazione di questi risultati è che, nonostante i punteggi di affidabilità e validità siano buoni, l'analisi del metodo con cui sono stati condotti gli studi inclusi ha evidenziato in alcuni casi delle carenze nella descrizione della procedura, infatti non è stato possibile soddisfare alcuni items delle checklist utilizzate poichè il criterio non veniva specificato nel testo. Ciò si riferisce in particolare alla modalità in cieco da parte degli esaminatori o dei soggetti relativamente ai risultati delle tecniche o a eventuali fattori di contesto. Ciò, oltre ad inficiare il punteggio del singolo studio, non permette un equo confronto tra le procedure d'esame.

Altra criticità è costituita dal fatto che la selezione del campione verte perlopiù su soggetti asintomatici, che quindi non rispecchiano la realtà in cui il terapista si trova a dover applicare le tecniche in oggetto.

Interessante è inoltre il dato emerso dallo studio sulla fiducia nella fissazione del segmento da parte dell'esaminatore: più è alto il livello di fiducia e più il livello fissato è preciso, il che aumenta proporzionalmente con il grado di concordanza tra gli esaminatori. Ciò in pratica clinica si traduce nel fatto che l'affidabilità delle tecniche è maggiore nel momento in cui queste sono praticate da operatori che posseggano un expertise tale da eseguirle con sicurezza e in modo standardizzato [32].

Un ulteriore limite che emerge dalla visione complessiva di questi risultati è che abbiamo tecniche di mobilità ben valutate in termini di affidabilità ma non vi è riscontro di validità in esse mentre si verifica la situazione opposta per quanto riguarda quelle di palpazione e individuazione dei livelli intervertebrali che, seppur abbiano un buon grado di validità, l'affidabilità non viene contemplata o comunque non raggiunge un livello sufficiente da poter essere vagliata.

La considerazione che questo lavoro ci permette di fare è che l' adozione di misure specificatamente selettive per l'inclusione degli studi ha determinato i medesimi risultati rispetto ad altre revisioni che avevano impostato parametri di selezione più ampi. Ne consegue che l'esecuzione di una revisione della letteratura inserendo una restrizione di criteri secondo data e metodologia di ricerca non cambia i risultati di validità e affidabilità.

Pertanto, oltre a essere in accordo a ciò che la letteratura precedente propone, i risultati di questo studio traslati nella quotidiana pratica clinica non fanno altro che confermare ciò che, seppur senza consistente evidenza scientifica di supporto, viene comunemente insegnato e constatato durante l'applicazione delle tecniche manuali in questione. Queste ultime in particolare risultano più accurate nel momento in cui vengono associate tra loro in quanto spesso si focalizzano su aspetti diversi di valutazione e proprio per questo è importante il processo che ne consegue di integrazione ed interpretazione di ciò che emerge dalla loro esecuzione. Il ragionamento clinico infatti, fondamentale per la formulazione della diagnosi funzionale e relativa stesura del piano di trattamento, deve prendere in considerazione la complessità dei dati che risultano dall'applicazione di tali metodiche, da correlare con quelli relativi agli step precedenti di valutazione non prettamente manuale; quest'ultima non è da sottovalutare in quanto, se ben ponderata, permette una selezione a priori dei test necessari, riducendo il numero di elementi da considerare che come evidenziato in alcuni studi costituirebbe un limite vista l'impossibilità di eseguire intere batterie di test in ogni situazione. Come già evidenziato, purtroppo questi risultati esulano a priori dalla nostra pratica clinica per l'aspetto relativo al campione in quanto è verosimilmente improbabile che i soggetti che si presenteranno alla nostra attenzione saranno asintomatici. Tuttavia a nostro favore vi è il dato secondo cui il punteggio della tecnica pragmatica risulta peggiore rispetto a quella standardizzata [32] e il che si traduce nel fatto che, nonostante le anomalie strutturali o variabilità anatomiche a cui è possibile trovarsi di fronte, la riproducibilità e la relativa efficacia della tecnica applicata saranno maggiori nel momento in cui questa seguirà determinati parametri, i medesimi che nel corso della sua pratica il terapista ha accuratamente sviluppato e fatto propri. Anch'esso quindi non giunge nuovo come concetto poiché è ormai risaputo che solo con la ripetuta e costante esecuzione della tecnica su un ampio campione di popolazione, il livello di affidabilità e accuratezza di essa incrementerà, proporzionalmente all'expertise del terapista appunto. Ciò permetterà l'aumento anche della fiducia nell'esecuzione, altro aspetto supportato dai dati descritti.

Nonostante questa revisione calzi adeguatamente a quelli che sono gli assiomi della pratica clinica in terapia manuale, preme sottolineare che i limiti evidenziati debbano essere tenuti in considerazione per i lavori futuri. Essi infatti dovranno curare i criteri di selezione del campione che dovrà essere maggiormente rappresentativo delle popolazione generale e analizzato specificando nel dettaglio le procedure d'esame utilizzate e, ultimo ma non per importanza, rispettare i criteri della valutazione metodologica utilizzata nelle precedenti revisioni al fine di poterne permettere il confronto.

6. CONCLUSIONI

La revisione condotta ha rilevato che i risultati osservati adottando un criterio restrittivo di selezione si sono dimostrati sovrapponibili ai dati raccolti dai precedenti lavori.

Indipendentemente dal livello di cut off della qualità metodologica, infatti, i risultati ottenuti denotano una scarsa accuratezza e riproducibilità delle tecniche manuali di identificazione dei punti di repere e valutazione della mobilità. Ciò si riscontra in particolar modo negli studi che testano la mobilità nel rachide lombare, la cui validità ed affidabilità risultano povere e migliorano moderatamente solo nel momento in cui ad essa si associa la provocazione del dolore e si valuta l'affidabilità intraoperatore. Risultati migliori si rilevano negli studi di palpazione di reperi, soprattutto per l'ambito dell'affidabilità, con innalzamento dei valori nel caso si utilizzi una combinazione di test di valutazione rispetto ad un singolo approccio, qualunque sia la tecnica scelta. Per quanto riguarda invece il rachide toracico si è riscontrata una maggiore affidabilità per la mobilità, specie in associazione con la provocazione del dolore, mentre la palpazione denota risultati più significativi per il criterio di validità sebbene l'accuratezza di alcuni aspetti della tecnica sia buona.

La mancanza di consistente validità in merito alle tecniche manuali utilizzate nella rilevazione di disfunzioni muscolo-scheletriche e la bassa affidabilità emersa nei test clinici costituiscono pertanto una sfida continua per la professione, sfida ormai che nella ricerca scientifica si deve allontanare da un focus costantemente mirato all'esclusiva identificazione di un punto di repere affidabile o alla quantificazione del movimento segmentario: risultati confermati da studi diversificati non necessitano di ulteriori ricerche.

Questo non significa che i dati poco incoraggianti in termini di validità e affidabilità delle procedure finora usate debbano portare all'eliminazione di tali tecniche manuali. Diventa invece importante cambiare il modello di ragionamento di fondo, il quesito clinico di cui tali test rappresentano lo strumento di risposta. Più che al riconoscimento del livello vertebrale corretto ed alla misurazione della sua mobilità infatti bisogna porre attenzione all'individuazione del/dei segmenti correlati ai disturbi del paziente ed alla capacità di individuarlo/i correttamente nei trattamenti successivi.

Oltre a ciò, i test manuali permettono al terapista di raccogliere informazioni multifattoriali, qualitative e quantitative, relative non solo al ROM, ma anche alla rigidità, al dolore, alle

risposte muscolari, psico-emotive e neurovegetative del paziente, e di rilevare le modificazioni di questi parametri sia intra che inter seduta.

L'integrazione e l'elaborazione di queste informazioni con l'esame anamnestico, all'interno quindi di una valutazione globale, consente pertanto di determinare la capacità di carico dei segmenti correlati al disturbo del paziente e scegliere in modo specifico la posologia di trattamento, adattandola in modo dinamico e coerente con la situazione clinica del singolo soggetto.

Un'ulteriore attenzione, ritornando pertanto ad una riflessione sulla direzione degli studi scientifici, può essere rivolta ad una ricerca, attraverso RTC, finalizzata all'outcome, ossia all'osservazione del cambiamento della prognosi del paziente in base all'applicazione o meno dei suddetti test diagnostici per rilevarne l'importanza in funzione dei risultati di trattamento. Non dobbiamo dimenticare infatti che un test diagnostico ha la sua ragione d'essere in termini di rilevanza nel momento in cui il suo utilizzo modifichi la prognosi del paziente stesso: una mera applicazione sulla base di sterili valori di affidabilità e validità ne screditerebbe l'essenza.

RINGRAZIAMENTI

La stesura della revisione è stata supportata dal lavoro in singolo di Chiara Dal Pos per la parte relativa a Risultati e Discussione del rachide lombare e da Giulia Franco per i capitoli Risultati e Discussione del rachide dorsale. Ringraziamo il relatore e terzo revisore, dott. Aldo Ciuro, per la sua efficiente assistenza e supervisione.

BIBLIOGRAFIA

1. Cooper, K., et al., *The use of pMRI to validate the identification of palpated bony landmarks.* Man Ther, 2013. **18**(4): p. 289-93.

2. Abbott JH, F.T., Fritz JM, Hing WA, Reid D, Whitman JM., *Manual physical assessment of spinal segmental motion: Intent and validity. .* Man Ther Nov 7, 2007.

3. Fritz JM, C.J., Childs JD. , *Subgrouping patients with low back pain: Evolution of a classification approach to physical therapy.* J Orthop Sports Phys Ther, 2007. **37:290-302.**

4. McGaugh JM, B.J., Dedrick GS, Jones EA, Sizer PS., *Comparing the anatomical consistency of the posterior superior iliac spine to the iliac crest as reference landmarks for the lumbopelvic spine: a retrospective radiological study.* Clin Anat., 2007 Oct. **20(7):819-25.**

5. Robinson, R., et al., *Reliability and validity of a palpation technique for identifying the spinous processes of C7 and L5.* Man Ther, 2009. **14**(4): p. 409-14.

6. Merz, O., et al., *Validity of palpation techniques for the identification of the spinous process L5.* Man Ther, 2013. **18**(4): p. 333-8.

7. Harlick, J.C., S. Milosavljevic, and P.D. Milburn, *Palpation identification of spinous processes in the lumbar spine.* Man Ther, 2007. **12**(1): p. 56-62.

8. Borghi, B., et al., *Soft tissue depression at the iliac crest prominence: a new landmark for identifying the L4-L5 interspace.* Minerva Anestesiol, 2012. **78**(12): p. 1348-56.

9. Kim, H.W., et al., *Interexaminer reliability and accuracy of posterior superior iliac spine and iliac crest palpation for spinal level estimations.* J Manipulative Physiol Ther, 2007. **30**(5): p. 386-9.

10. Hestbaek L, L.-Y.C., *Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review.* J Manipulative Physiol Ther. , 2000 May. **23(4):258-75.**

11. Haneline, M.T. and M. Young, *A review of intraexaminer and interexaminer reliability of static spinal palpation: a literature synthesis.* J Manipulative Physiol Ther, 2009. **32**(5): p. 379-86.

12. Johansson, F., *Interexaminer reliability of lumbar segmental mobility tests.* Man Ther, 2006. **11**(4): p. 331-6.

13. Chakraverty, R.C., et al., *Identification of the correct lumbar level using passive intersegmental motion testing.* Anaesthesia, 2007. **62**(11): p. 1121-5.

14. Landel, R., et al., *Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing.* Phys Ther, 2008. **88**(1): p. 43-9.

15. Troke, M., D. Schuit, and C.M. Petersen, *Reliability of lumbar spinal palpation, range of motion, and determination of position.* BMC Musculoskelet Disord, 2007. **8**: p. 103.

16. Schneider, M., et al., *Spinal palpation for lumbar segmental mobility and pain provocation: an interexaminer reliability study.* J Manipulative Physiol Ther, 2008. **31**(6): p. 465-73.

17. Hart J, N.C., *Allowing a possible margin of error when assessing student skills in spinous process location.* J Chiropr Educ., 2011 Fall. **25(2):182-5.**

18. Snider, K.T., et al., *Palpatory accuracy of lumbar spinous processes using multiple bony landmarks.* J Manipulative Physiol Ther, 2011. **34**(5): p. 306-13.

19. Deore, M. and S. May, *The inter-rater and intra-rater reliability of passive physiological accessory movement assessment of lumbar spine in novice manual therapists.* J Bodyw Mov Ther, 2012. **16**(3): p. 289-93.

20. Kilby, J., N.R. Heneghan, and M. Maybury, *Manual palpation of lumbo-pelvic landmarks: a validity study.* Man Ther, 2012. **17**(3): p. 259-62.

21. Shaw, K.A., et al., *Establishing the content validity of palpatory examination for the assessment of the lumbar spine using ultrasonography: a pilot study.* J Am Osteopath Assoc, 2012. **112**(12): p. 775-82.

22. Ernst MJ, R.F., Bauer CM, Marcar VL, Kool J., *Determination of thoracic and lumbar spinal processes by their percentage position between C7 and the PSIS level.* BMC Res Notes., 2013 Feb 11. **6:58.** .

23. Hidalgo, B., et al., *Intertester agreement and validity of identifying lumbar pain provocative movement patterns using active and passive accessory movement tests.* J Manipulative Physiol Ther, 2014. **37**(2): p. 105-15.

24. Bracht, M.A., et al., *Inter- and Intra-observer Agreement of the Motion Palpation Test for Lumbar Vertebral Rotational Asymmetry.* Physiother Can, 2015. **67**(2): p. 169-73.

25. Phillips, D.R., et al., *Simple anatomical information improves the accuracy of locating specific spinous processes during manual examination of the low back.* Man Ther, 2009. **14**(3): p. 346-50.

26. Atya, A.M., *The validity of spinal mobility for prediction of functional disability in male patients with low back pain.* J Adv Res, 2013. **4**(1): p. 43-9.

27. Carlesso, L.C., et al., *Beliefs and practice patterns in spinal manipulation and spinal motion palpation reported by canadian manipulative physiotherapists.* Physiother Can, 2013. **65**(2): p. 167-75.

28. Lakhani, E., et al., *Motion palpation used as a postmanipulation assessment tool for monitoring end-feel improvement: a randomized controlled trial of test responsiveness.* J Manipulative Physiol Ther, 2009. **32**(7): p. 549-55.

29. Stochkendahl, M.J., et al., *Manual examination of the spine: a systematic critical literature review of reproducibility.* J Manipulative Physiol Ther, 2006. **29**(6): p. 475-85, 485.e1-10.

30. Brismee, J.M., et al., *Interrater reliability of a passive physiological intervertebral motion test in the mid-thoracic spine*. J Manipulative Physiol Ther, 2006. 29(5): p. 368-73.

31. Heiderscheit, B. and W. Boissonnault, *Reliability of joint mobility and pain assessment of the thoracic spine and rib cage in asymptomatic individuals*. J Man Manip Ther, 2008. 16(4): p. 210-6.

32. Walker, B.F., et al., *Interrater Reliability of Motion Palpation in the Thoracic Spine.* Evid Based Complement Alternat Med, 2015. 2015: p. 815407.

33. Cooperstein R, Haneline M., *Spinous process palpation using the scapular tip as a landmark vs radiographic criterion standard.* J Chiropr Med, 2007. 6(3): 87–93.

34. Cooperstein R, Young M. *Mapping intended spinal site of care from the upright to prone position: an interexaminer reliability study*. Chiropr Man Therap., 2014. 16;22:20.

35. Ernst MJ, Rast FM, Bauer CM, Marcar VL, Kool J. D*etermination of thoracic and lumbar spinal processes by their percentage position between C7 and the PSIS level.* BMC Res Notes. 2013 Feb 11;6:58. doi: 10.1186/1756-0500-6-58.

36. Cooperstein, R., M. Haneline, and M. Young, *Interexaminer reliability of thoracic motion palpation using confidence ratings and continuous analysis.* J Chiropr Med, 2010. 9(3): p. 99-106.

37. Geelhoed, M.A., et al., *A new model to facilitate palpation of the level of the transverse processes of the thoracic spine.* J Orthop Sports Phys Ther, 2006. 36(11): p. 876-81.

38. Holmgren U., Waling K., *Inter-examiner reliability of four static palpation tests used for assessing pelvic dysfunction.* Manual Therapy 13, 2008. 50–56.

39. Sutton C., Nono L., Johnston R.G., Thomson O.P., *The effects of experience on the inter-reliability of osteopaths to detect changes in posterior superior iliac spine levels using a hidden heel wedge.* Journal of Bodywork & Movement Therapies, 2013. 17, 143e150.