UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Programa Isquios: estudio de la salud de la espalda a través de redes de Inteligencia artificial en escolares de 6 a 16 años

Dña. María Teresa Martínez Romero
2021

UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

PROGRAMA ISQUIOS: ESTUDIO DE LA SALUD DE LA ESPALDA A TRAVÉS DE REDES DE INTELIGENCIA ARTIFICIAL EN ESCOLARES DE 6 A 16 AÑOS

Dª. María Teresa Martínez Romero

Directores:
María Pilar Sainz de Baranda Andújar
Fernando Santonja Medina
2021

Esta tesis ha sido financiada a través de un contrato predoctoral del Ministerio de Educación, Cultura y Deporte (identificación: FPU15/05200). Esta tesis es parte del Proyecto titulado "La prevención del dolor de espalda a través de la Educación Física" (DEP2010-21793), financiado por Ministerio de Ciencia e Innovación en el marco del subprograma de proyectos de investigación fundamental no orientada en I+D+i (2010-2013). Parte de esta tesis fue realizada durante una estancia de investigación en la Facultad de Ciencias del Deporte y el Ejercicio en la Universidad de Gloucestershire (Reino Unido) desde 01/05/2018 al 01/08/2018 que fue financiada por el Ministerio de Educación, Cultura y Deporte.

This thesis was funded by a predoctoral grant awarded by the Ministerio de Educación, Cultura y Deporte (grant number: FPU15/05200). This study is part of the project entitled "La prevención del dolor de espalda a través de la Educación Física" (DEP2010-21793). Part of this thesis was carried out during the research stay at the School of Sport and Exercise of the University of Gloucestershire from 01/05/2018 to 01/08/2018 funded by the Ministerio de Educación, Cultura y Deporte.

UNIVERSIDAD DE MURCIA
DEPARTAMENTO ACTIVIDAD FÍSICA Y DEPORTE
Facultad de Ciencias del Deporte

María del Pilar Sainz de Baranda Andújar

Doctora por la Universidad de Murcia y Profesora Titular del Departamento de Actividad Física y Deporte de la Universidad de Murcia

AUTORIZA:

La presentación de la tesis doctoral titulada: Programa ISQUIOS: Estudio de la salud de la espalda a través de redes de inteligencia artificial en escolares de 6 a 16 años, realizada por Dª. María Teresa Martínez Romero, bajo mi inmediata dirección y supervisión, y que presenta para la obtención del Grado de Doctor por la Universidad de Murcia.

Y, para que surta los efectos oportunos al interesado, firmo la presente en Murcia, a nueve de diciembre de dos mil veinte.

Dª. María del Pilar Sainz de Baranda Andújar

UNIVERSIDAD DE MURCIA
DEPARTAMENTO DE CIRUGÍA, PEDIATRÍA, OBSTETRICIA Y GINECOLOGÍA
Facultad de Medicina

Fernando Santonja Medina

Doctor en Medicina y Cirugía y Profesor Titular del
Departamento de Cirugía, Pediatría, Obstetricia y Ginecología de la
Universidad de Murcia

AUTORIZA:

La presentación de la tesis doctoral titulada: Programa ISQUIOS: Estudio de la salud de la espalda a través de redes de inteligencia artificial en escolares de 6 a 16 años, realizada por Dª. María Teresa Martínez Romero, bajo mi inmediata dirección y supervisión, y que presenta para la obtención del Grado de Doctor por la Universidad de Murcia.

Y, para que surta los efectos oportunos al interesado, firmo la presente en Murcia, a nueve de diciembre de dos mil veinte.

D. Fernando Santonja Medina

Dedicado a mi abuela Juanita.
Sigue cuidándome desde allí arriba.

"Si te atreves a enseñar,
no dejes de aprender" (John Cotton Dana)

AGRADECIMIENTOS

Siempre he pensado que una persona con apenas 18 años no tiene la capacidad ni la experiencia para poder decidir qué estudiar o en qué trabajar. Aunque exista vocación, la vida da demasiadas vueltas y nunca sabes qué puede pasar. Cuando yo tenía 18 años no tenía nada claro, medicina, enfermería, fisioterapia, CAFD...y 10 años después aquí estoy. Ahora puedo decir que elegí bien y sé hacia dónde quiero ir a partir de ahora. Y esa sensación, me encanta.

Por eso quiero dar las gracias a mi familia, no sólo por haberme apoyado de forma incondicional de principio a fin, sino porque sin vosotros hoy no sería quien soy ahora ni hubiera podido conseguir todo lo que me he propuesto y lo que vendrá. Gracias mamá, papá, Victoria y Sandra por estar siempre.

Estos 10 años de andadura universitaria no hubiesen sido lo mismo sin ti, Fran. ¿Quién te iba a decir que al mes de empezar la carrera todos tus planes iban a cambiar por mi culpa? Aunque por suerte para los dos, los cambios siempre fueron a mejor y eso es algo que yo tenía clarísimo. Hemos crecido personal y profesionalmente de la mano, ayudándonos y haciéndonos mejor el uno al otro. Me siento muy afortunada de haberte encontrado. Gracias porque contigo soy mejor. Y gracias a tu familia, porque sin serlo, me siento como una hija o hermana más.

También quiero dar las gracias a mis amigas porque a pesar de la distancia o del tiempo que pasemos sin vernos, todo es igual, nada cambia. Y eso, después de más de 20 años juntas, es muy difícil de conseguir.

Por último, y no menos importante, quiero expandir mis agradecimientos a todas las personas que forman parte del Grupo de Investigación Aparato Locomotor y Deporte, porque hemos compartido muchos momentos juntos en clases, congresos, mediciones, cenas, fiestas. Todo ello ha hecho que seamos, más que compañeros, amigos.

Gracias a mi directora, Pilar Sainz de Baranda, porque tú me descubriste el mundo de la investigación. Cuando empecé la carrera mi idea era ser profesora de Educación Física, pero cuando te conocí en el último año todo cambió. Desde ese momento supe que quería dedicarme a investigar y a ser docente en la universidad. Gracias a ti conseguí el contrato predoctoral que me permitió hacer esta tesis y formarme como docente. Además, tanto el Programa ISQUIOS como Vertebrín© formarán parte de mi vida para siempre. En estos años he aprendido muchísimo y eso ha sido gracias a ti y a Fernando Santonja, mi visión sobre la columna vertebral y su comportamiento son totalmente diferentes a cuando empecé, ahora entiendo el "Morfotipo Sagital Integral".

Gracias también a mi familia de "becarios" Fran Robles, Alba Aparicio y Mª José García porque sin vuestra compañía, estos años de tesis se habrían hecho muy cuesta arriba. La universidad nos ha llevado a forjar una gran amistad porque entre "becarios" nos entendemos y comprendemos a la perfección. Alba, aunque termine la tesis, siempre podrás contar con mi apoyo en "Cheltenham Crew".

Antonio Cejudo, Raquel Hernández, Olga Rodríguez y Francisco Ayala, gracias por dejarme aprender a vuestro lado y de vosotros. Sois unos grandísimos docentes, compañeros y amigos, aportáis mucho tanto a los estudiantes como a la sociedad. Me encantaría seguir creciendo y mejorando a vuestro lado. Personas como vosotras las quiero siempre en mi equipo.

No me puedo olvidar de mis compañeros del 5-10 y de las miles de mediciones que hemos hecho juntos. Muchas gracias Alfonso, Chema, Fernando, Andrea y Angélica, sois unos compañeros increíbles y unas grandísimas personas. Tengo mucha suerte al haber podido compartir esta etapa de mi vida junto a vosotros.

Por último, agradecer la financiación recibida por el Ministerio de Educación, Cultura y Deporte a través de un contrato predoctoral (FPU15/05200) para la realización de esta tesis doctoral.

Tabla de contenidos / Table of contents

Lista de Figuras / List of Figures

Lista de Tablas / List of Tables

Lista de Anexos / List of Appendix

Abreviaturas / Abbreviations

ADL: activities of daily life
ANOVA: analysis of variance
BMI: body mass index
BP: back pain
BF: Bayesian factor
BFP: bending forward position
BS: Biering-Sorensen
BTC: Bench Trunk Curl-Up
CA: chronological age
CCI: coeficiente de correlación intraclase
CI: confidence interval
CI: credible interval
CU: Curl-Up
CV: coefficient of variation
DE: dolor espalda
DEE: Dynamic Extensor Endurance
DL: dolor lumbar
EC: edad cronológica
EM: edad madurativa
EVA: escala visual analogica
FE: Flexor Endurance
GBD: Global burden disease
ICC: intraclass correlation coefficient
IMC: índice de masa corporal
ITF: Isometric Trunk Flexion endurance
LBP: low back pain
L-H: lumbo-horizontal
L-V: Lumbo-vertical
MBP: mid-back pain
MDC: minimum detectable change
NP: neck pain
OMS: Organización Mundial de la Salud
OR: Odd ratio
PCU: Partial Curl-Up
PDSR: Prone Double Straight-leg Raise
PE: physical education
PHE: passive hip extension
PHF_KE: passive hip flexion with knee extended
PHV: peak height velocity
PIH: Plank Isometric Hold
PICR: prone isometric chest raise
PVC: pico de velocidad de crecimiento
RG: reliability generalization
RI: reliability induction
ROM: range of motion
SB: side bridge
SD: standard deviation
SP: standing position
SSP: slump sitting position
SWC: small worthwhile change
TE: typical error
TEM: typical error of measurement
VAS: visual analogue scale

Resumen

Los objetivos de los estudios que constituyen la presente tesis fueron: a) estudiar la prevalencia del dolor de espalda (DE) y sus características y determinar los factores físicos asociados con el DE en niños y adolescentes de la Región de Murcia, a través del protocolo "Fitness Postural", y b) conocer las pruebas de campo que se emplean para la valoración de la resistencia muscular del tronco, determinar su fiabilidad y analizar la resistencia de los músculos del tronco en niños y adolescentes de la Región de Murcia en función de su estado madurativo.

Estudio 1: Los objetivos de este estudio fueron: a) describir la prevalencia de escolares de la Región de Murcia (España) que sufrieron DE en el año o mes anterior en cuanto a sexo, edad cronológica (EC), etapa de maduración, índice de masa corporal (IMC) y las características del dolor (zona, frecuencia, intensidad y severidad), y b) determinar los posibles factores asociados con el DE. Se utilizó una muestra de 513 estudiantes, con edades comprendidas entre los 9 y los 16 años (12,6 ± 1,9 años); 257 (50,1%) eran chicos y 256 (49,9%) chicas. La presencia de DE en el último año se observó en 180 escolares (35,1%), 81 eran chicos (45%) y 99 eran chicas (55%), no se encontró asociación entre el DE y el sexo. Cuando se analizó la prevalencia de 1 mes, 89 (17,3%) escolares presentaron DE, de los cuales 36 eran chicos (40,5%) y 53 chicas (59,5%), encontrándose asociación entre tener DE el mes anterior y el sexo femenino. Según la edad cronológica de los participantes, la prevalencia de DE en el último año y el último mes creció a medida que aumentaba la edad de los estudiantes. La misma tendencia apareció al observar la prevalencia de DE en el último año y el último mes según el estado madurativo. De acuerdo con la clasificación del IMC, la prevalencia de DE también aumentó desde aquellos escolares con bajo o normopeso a aquellos escolares con sobrepeso u obesidad. En los niños, se encontraron asociaciones entre el grupo de EC y DE en el último año y mes. También se encontró una asociación entre tener DE durante el último año y el IMC. Por otro lado, en las niñas se encontraron asociaciones entre el DE en el último año y el último mes y los grupos de EC, así como con la etapa de maduración. La zona de la columna donde se señaló dolor con mayor frecuencia fue la zona lumbar (59,4%), seguida de la zona torácica de la espalda (37,2%) y la región cervical (21,1%). Casi la mitad de los participantes (48,9%) respondieron haber tenido DE "algunos días" en el último año. De los 180 estudiantes con DE, 22 (12,3%) experimentaron dolor que les impedía realizar actividades de la vida diaria y 11 (6,1%) indicaron padecer ciática. La intensidad media de DE fue de 3,87 ± 2,01 en la escala visual analógica. En resumen, el presente estudio

mostró una prevalencia de DE asociada al estado de maduración y al peso de los participantes, encontrándose diferentes patrones de prevalencia según el sexo.

Estudio 2: Los objetivos del presente estudio fueron a) estudiar las características antropométricas y el "Fitness Postural" en niños y adolescentes con y sin DE, según el sexo, y b) determinar si estos factores físicos son posibles desencadenantes del DE. El protocolo de valoración de " Fitness Postural" se basa en la medición de los componentes biomecánicos que influyen en la salud de la espalda mediante la evaluación de las curvaturas sagitales de la columna vertebral, la inclinación pélvica, el rango de movimiento de la cadera (ROM) y la resistencia de los músculos del tronco. Se midieron 252 estudiantes, con un rango de edad entre los 9 y los 15 años. En la valoración de la curvatura torácica los chicos presentaron una mayor cifosis en todas las posiciones evaluadas. Por el contrario, para la curva lumbar, fueron las niñas las que presentaron mayor lordosis en bipedestación, mientras que en sedentación y flexión del tronco presentaron una cifosis lumbar menos pronunciada, junto con un mayor porcentaje de pelvis equilibradas. En cuanto al ROM, las niñas mostraron un rango mayor en todas las pruebas realizadas (ROM de los flexores y extensores de la cadera). Finalmente, en las pruebas de resistencia de los músculos del tronco, los niños mostraron un mayor rendimiento en los músculos flexores-laterales y las niñas en la musculatura extensora. La presencia de DE en el último año, se observó en 71 (28,2%) escolares, de los cuales 35 eran niños (28,7%) y 36 eran niñas (27,7%). Los participantes con DE mostraron mayores valores en las variables relacionadas con el peso corporal, encontrando estas diferencias también entre los chicos. En la valoración de la columna, solo se observaron diferencias significativas entre niñas con y sin DE para la curvatura lumbar en flexión máxima del tronco, encontrando menor cifosis lumbar en las niñas con DE. Los participantes con DE tenían un ROM en los flexores de cadera más reducido, estas diferencias también se encontraron en las niñas y en los niños se observó una tendencia hacia la significancia ($p = 0.06$). Por último, en las pruebas de resistencia muscular del tronco, solo se encontraron diferencias significativas entre chicas con y sin DE para la prueba Ito. En resumen, el presente estudio mostró que los niños y adolescentes pre-púberes y durante el estirón puberal presentaron diferencias relacionadas con el sexo en la evaluación del protocolo de "Fitness Postural". En cuanto al DE, se encontró asociación entre el DE en los niños y el IMC más alto, y entre el DE en las niñas y el ROM reducido en los músculos flexores de la cadera y la baja resistencia de los flexores del tronco.

Estudio 3: El objetivo del presente meta-análisis fue estimar la fiabilidad inter e intra-examinador de las medidas de resistencia obtenidas a través de pruebas de campo de extensión del tronco y explorar la influencia de los moderadores en las estimaciones de fiabilidad. También se calculó la tasa de inducción de fiabilidad de las medidas de resistencia de extensión del tronco. Se realizó una búsqueda sistemática utilizando varias bases de datos y, posteriormente, se seleccionaron 28 estudios que informaron de coeficientes de correlación intraclase (CCI) para las medidas de resistencia de extensión del tronco. Se realizaron meta-análisis separados, utilizando un modelo de efectos aleatorios. Cuando fue posible, se llevaron a cabo análisis de posibles variables moderadoras. El valor de fiabilidad promedio (CCI) inter-examinador para la medida de resistencia obtenida de la prueba de Biering-Sorensen fue de 0,94. Las estimaciones de fiabilidad intra-sesión (CCI) de las medidas de resistencia registradas mediante la prueba de Biering-Sorensen, la prueba Prone Isometric Chest Raise y la prueba Prone Double Straight-leg Raise fueron 0,88, 0,90 y 0,86, respectivamente. La fiabilidad promedio entre sesiones (CCI) de las medidas de resistencia de la prueba de Biering-Sorensen, la prueba Prone Isometric Chest Raise y la prueba Dynamic Extensor Endurance fueron 0,88, 0,95 y 0,99, respectivamente. Sin embargo, debido a la evidencia limitada disponible, las estimaciones de fiabilidad de las medidas obtenidas a través de las pruebas Prone Isometric Chest Raise, Prone Double Straight-leg Raise y Dynmaic Extensor Endurance deben considerarse con cierto grado de precaución. Los instrumentos de control de posición, el material y la sesión de familiarización demostraron una asociación estadística con la fiabilidad entre sesiones de la prueba de Biering-Sorensen. La tasa de inducción de la fiabilidad fue del 72,8%. Solo la medida de resistencia de extensión del tronco obtenida a través de la prueba de Biering-Sorensen, presentó evidencia científica suficiente en términos de fiabilidad para justificar su uso con fines prácticos y de investigación.

Estudio 4: Este estudio tuvo como objetivo explorar la fiabilidad entre sesiones de las medidas obtenidas de 2 pruebas de resistencia para la extensión de tronco (pruebas de Biering-Sorensen y Dynamic Extensor Endurance (DEE)) y 3 para la flexión de tronco (Ito, Side Bridge y Bench Trunk Curl-Up (Pruebas BTC)) en adolescentes de 4 cursos diferentes de Educación Secundaria Obligatoria. Un total de 208 (chicos, n = 95; chicas, n = 113) adolescentes (de 12 a 18 años) realizaron todas las pruebas de campo en 2 sesiones separadas, con 7 días de diferencia. Se calcularon los valores de fiabilidad para todos los participantes agrupados en el mismo conjunto de datos y por separado según la edad, mediante la fiabilidad relativa (coeficiente de correlación intraclase (CCI)), las diferencias entre sesiones (sesgo sistemático) y la precisión de las mediciones, es decir, fiabilidad absoluta (error típico expresado como coeficiente de variación (CV_{TE}) y mínimo cambio detectable (MDC_{95})). No se encontraron diferencias relevantes relacionadas con los cursos para la fiabilidad o la sensibilidad test-retest en cada prueba, por lo que los valores agrupados se consideraron generalizables para esta cohorte de adolescentes. La mayoría de las medidas de resistencia del tronco demostraron una fiabilidad relativa aceptable (los CCI oscilaron entre 0,75 y 0,94). Sin embargo, se identificaron diferencias significativas entre sesiones para las medidas de las pruebas DEE y BTC. Asimismo, la precisión de la medición de cada prueba de campo fue pobre (CV_{TE} osciló entre 11,3 y 29,4%) y el MDC_{95} reveló que son necesarios cambios del 42%, para las pruebas de resistencia de extensión del tronco y del 31,4%, para las pruebas de resistencia de flexión del tronco después de una intervención para indicar un cambio significativo por encima del error de medición. Todas las pruebas fueron lo suficientemente sensibles como para detectar cambios de moderados a grandes en la resistencia de los músculos del tronco. Los hallazgos de este estudio indican que solo la prueba BTC demuestra una fiabilidad aceptable entre sesiones (ICC> 0,9, CVTE ~ 10%, MDC95 ~ 30%) para monitorear los cambios en los valores de resistencia del tronco que se pueden esperar en adolescentes después de realizar un programa de intervención. El uso de sesiones de familiarización supervisadas antes de realizar las pruebas y un fuerte estímulo para realizar un esfuerzo máximo en cada prueba, pueden ser estrategias útiles para mejorar los valores de fiabilidad.

Estudio 5: Este estudio tuvo como objetivo analizar y comparar la influencia de la edad cronológica (EC) y la edad madurativa (EM) en la resistencia de los músculos del tronco en chicos y chicas en edad escolar. Se utilizaron 2 pruebas de resistencia para la extensión de tronco (pruebas de Biering-Sorensen (BS) y Dynamic Extensor Endurance (DEE)) y 3 para la flexión de tronco (Ito, Side Bridge (SB) y Bench Trunk Curl-Up (BTC)). El análisis entre grupos mediante ANOVA bayesiano mostró diferencias, con al menos evidencia sólida, a favor de la hipótesis alternativa (H_1) con un tamaño de efecto al menos moderado ($\delta > 0,6$) para las puntuaciones de resistencia de todas las pruebas del tronco realizadas y la EC, independientemente del sexo, es decir, todas las medidas de resistencia del tronco aumentaron con la edad. Dentro de los niños, se identificaron diferencias de EC para todas las pruebas de campo de resistencia del tronco. Entre las niñas, se encontraron diferencias de EC para todas las pruebas de campo de resistencia del tronco, excepto para la prueba Dynamic Extensor Endurance (DEE). En ambos casos, las diferencias se encontraron principalmente entre los participantes más jóvenes, con peor rendimiento y los de mayor edad, con mayor resistencia. Se observaron diferencias en función del sexo y EC para la prueba DEE entre chicos y chicas de 14 años y entre los mayores de 15 años. También se encontraron diferencias relacionadas con el sexo en la prueba SB para el lado derecho e izquierdo entre las participantes mayores de 15 años. El análisis de ANOVA bayesiano mostró una fuerte evidencia a favor del H1 entre las etapas de maduración, independientemente del sexo, para todas las pruebas de resistencia del tronco. Entre los niños, se encontraron diferencias antes y después del pico de velocidad de crecimiento (PVC) para todas las pruebas de campo de resistencia del tronco. Solo se encontraron diferencias entre pre y circa-PVC en hombres para las pruebas BTC, Ito, DEE y SB. En las niñas, se encontraron diferencias antes y después de la PVC para todas las pruebas de campo de resistencia del tronco, excepto para las pruebas DEE y SB. Solo se encontraron diferencias entre circa- y post-PVC en mujeres para las pruebas BTC e Ito. Para todas las pruebas de campo de resistencia del tronco, los hombres presentaron un rendimiento más alto que las mujeres en circa-PVC, aunque para las pruebas BS y SB los resultados no son relevantes. También se encontraron diferencias relacionadas con el sexo para las pruebas DEE y SB entre niños y niñas después del brote de crecimiento. En conclusión, los niños y niñas pre-púberes, mostraron un desempeño similar en todas las pruebas de campo, cuando entraron en el estirón puberal las diferencias fueron relevantes (BTC, Ito, DEE), con los niños mostrando puntuaciones más altas que las niñas en todas las pruebas. Después del PVC, las diferencias se mantuvieron para los flexores laterales del tronco (prueba SB) y la prueba DEE y se igualaron para las pruebas de flexores del tronco (pruebas BTC e Ito) y la prueba BS (Biering-Sorensen).

Abstract

The aims of the studies that set up this thesis were: a) to study the prevalence of back pain (BP) and its characteristics and determine the physical factors associated with BP in children and adolescents from the Region of Murcia through the protocol "Postural Fitness", and b) know the field-based tests that are used to assess the trunk muscular endurance, determine its reliability and analyse the trunk muscle endurance in children and adolescents from the Region of Murcia according to their maturational state.

Study 1: The aims of this study were a) to describe the prevalence of school-aged children from the Region of Murcia (Spain) who suffered from BP within the previous year or month in terms of sex, CA, maturational stage, BMI, affected region of the back, frequency, intensity, and severity of pain, and b) to determine the possible factors associated with this disorder. This study was based on a sample of 513 students, aged between 9 and 16 years (12.6±1.9 years); 257 (50.1%) were males and 256 (49.9%) females. The presence of BP in the last year was observed in 180 (35.1%) students, 81 were boys (45%) and 99 were girls (55%), no association was found between BP and sex. For the 1-month prevalence, 89 (17.3%) students presented BP, of which 36 were boys (40.5%) and 53 were girls (59.5%), finding an association between having BP the previous month and being a girl. According to the CA of the participants, the prevalence of BP in the last year and last month grew as the students' age increased. The same trend appeared when observing the prevalence of BP in the last year and last month according to the maturational stage. According to the BMI classification, the prevalence of BP also increased from those with low or normal weight to those with overweight or obesity, finding an association between being overweight or obese and presenting BP in the last year. In boys, there were associations between the CA group and BP in the last year and month. An association was also found between having BP during the past year and BMI. On the other hand, in girls, associations were found between BP in the last year and the last month and the CA groups, as well as with the maturation stage. The spinal area where the pain was most often reported was the lower back (59.4%), followed by the mid-back (37.2%) and neck area (21.1%). Nearly half of participants (48.9%) reported having BP "a few days" in the past year. Of the 180 students with BP, 22 (12.3%) experienced pain that prevented them from performing activities of daily living (ADL) and 11 (6.1%) indicated suffering sciatica. The mean intensity of BP reported was 3.87±2.01 on the visual analogue scale. In summary, the present study showed a prevalence of BP

associated with the maturational state and the weight of participants, finding different prevalence patterns after adjusted by sex.

Study 2: The aims of the present study were a) to study the anthropometric characteristics and "Postural Fitness" in children and adolescents with and without BP according to sex, and b) to determine if these physical factors are possible triggers of BP. The "Postural Fitness" protocol is based on the measurement of the biomechanical components that influence on the health of the back through the assessment of the sagittal spinal curvatures and pelvic tilt, the hip range of motion (ROM) and the trunk muscles endurance. This study was based on a sample of 252 students, aged from 9 to 15 years. In the assessment of thoracic spinal curvature boys presented a greater kyphosis in all the assessed positions. Conversely, for the lumbar curve, it was the girls who showed greater lordosis in standing position and lower lumbar kyphosis, together with greater neutral pelvis in slump sitting and in bending forward positions. About ROM, girls showed a greater range in all the tests performed (flexors and extensors ROM). Finally, in the trunk muscle endurance tests, boys showed greater performance in the lateral-flexor muscles and girls in the extensor musculature. Regarding the pelvic tilt assessment, most of the students presented posterior pelvic tilt, especially when they bending forward. A majority of participants had a reduced ROM in the hip extensors and flexors, being much more prevalent among boys. The presence of BP in the last year was observed in 71 (28.2%) students, of students with BP in the last year, 35 were boys (28.7%) and 36 were girls (27.7%). Participants with BP showed an increase in the variables related to body weight, finding these differences were also among boys. In the spine assessment, only significant differences were observed between girls with and without BP for lumbar curvature in bending forward position, finding less lumbar kyphosis in girls with BP. Participants with BP had a lower ROM, these differences were also found in girls and in boys, a trend towards significance was observed (p=0.06). Lastly, in the trunk muscle endurance tests, only significant differences were found between girls with and without BP for the Ito test. Associations were found between a reduced ROM and BP. In summary, the present study showed that children and adolescents in pre- and circa-PHV presented sex-related differences in the assessment of "Postural Fitness" protocol. Concerning BP, association were found between BP in boys and higher IMC, and between BP in girls and reduced ROM in hip flexor muscles and poor trunk flexor endurance.

Study 3: This meta-analysis aimed to estimate the inter- and intra-tester reliability of endurance measures obtained through trunk extension field-based tests and to explore the influence of the moderators on the reliability estimates. The reliability induction rate of trunk extension endurance measures was also calculated. A systematic search was conducted using various databases, and subsequently 28 studies were selected that reported intraclass correlation coefficients for trunk extension endurance measures. Separate meta-analyses were conducted using a random-effects model. When possible, analyses of potential moderator variables were carried out. The inter-tester average reliability of the endurance measure obtained from the Biering-Sorensen test was intraclass correlation coefficient (ICC) = 0.94. The intra-session reliability estimates of the endurance measures recorded using the Biering-Sorensen test, the prone isometric chest raise test, and the prone double straight-leg test were ICC = 0.88, 0.90, and 0.86, respectively. The inter-session average reliability of the endurance measures from the Biering-Sorensen test, the prone isometric chest raise test, and the dynamic extensor endurance test were ICC = 0.88, 0.95, and 0.99, respectively. However, due to the limited evidence available, the reliability estimates of the measures obtained through the prone isometric chest raise, prone double straight-leg, and dynamic extensor endurance tests should be considered with a degree of caution. Position control instruments, tools, and familiarization session demonstrated a statistical association with the inter-session reliability of the Biering-Sorensen test. The reliability induction rate was 72.8%. Only the trunk extension endurance measure obtained through the Biering-Sorensen test presented sufficient scientific evidence in terms of reliability to justify its use for research and practical purposes.

Study 4: This study aimed to explore the inter-session reliability of the measures obtained from 2 trunk extension (Biering-Sorensen and Dynamic Extensor Endurance (DEE) tests) and 3 trunk flexion (Ito, Side Bridge and Bench Trunk Curl-Up (BTC) tests) endurance field-based tests in adolescents of 4 different high school age-based grades. A total of 208 (males, n = 95; females, n = 113) adolescents (ranging from 12 to 18 years) performed all the field-based tests on 2 separate testing sessions, 7-days apart. Both grouped (all participants pooled in the same data set) and separately by age-based grade inter-session reliability scores were explored through relative reliability (intraclass correlation coefficient (ICC)), inter-session differences (systematic bias) and precision of measurements (i.e. absolute reliability) (standard error of measurement expressed as a percentage of the mean score (CV_{TE}) and minimal detectable change (MDC_{95})). The sensitivity of each test was also assessed through the smallest worthwhile percentage change (SWC). No relevant age grade-related differences were found for either test-retest reliability or sensitivity in each test, so the grouped scores were considered as generalizable for this cohort of high school-aged adolescents. Most of the trunk endurance measures demonstrated acceptable relative reliability (ICCs ranged from 0.75 to 0.94). However, significant inter-session differences were identified for measures from the DEE and BTC tests. Likewise, the precision of the measurement of each field-based test was poor (CVTE ranged from 11.3 to 29.4%) with the MDC_{95} revealing that changes higher than 42% for trunk extension endurance tests and 31.4% for trunk flexion endurance tests after an intervention are required to indicate a significant change above measurement error. All tests were sensitive enough to detect moderate to large changes in trunk muscle endurance. Therefore, the findings from this study indicate that only the BTC test demonstrates acceptable inter-session reliability (ICC > 0.9, CV_{TE} ~ 10%, MDC_{95} ~ 30%) to monitor the changes in trunk endurance scores that may be expected in adolescents after performing an intervention program. The use of supervised familiarization sessions before performing the tests and strong encouragement to perform a maximal effort in each test may be helpful strategies to improve the reliability scores.

Study 5: This study aimed to analyse and compare the influence of CA and maturational stage on trunk muscles endurance in males and females school-aged. The between-group analysis through Bayesian ANOVA showed differences with at least strong evidence in favour of the alternative hypothesis (H1) with at least moderate effect size (δ >0.6) for the endurance scores of all trunk measures obtained from BTC, Ito, DEE, BS, SB-R and SB-L tests and CA groups, regardless of sex, i.e. all trunk endurance measures increased with rising age. Within boys, CA differences were identified for all trunk endurance field-based tests. Among girls, CA differences were found for all trunk endurance field-based tests, except for the DEE test. In both cases, the differences were mainly found between the youngest, worst performance, and oldest participants, highest endurance. Differences by sex and CA group were identified for the DEE test between males and females of 14 years and between those older than 15 years. Sex-related differences were also found in the SB test for the right and left side between participants older than 15 years. The Bayesian ANOVA analysis showed strong evidence in favour of the H1 between maturational stages, regardless of sex, for all trunk endurance values. Among boys, pre- and post-PHV differences were found for all trunk endurance field-based tests. Differences between pre- and circa-PHV were only found in males for BTC, Ito, DEE and SB-R tests. Within girls, pre- and post-PHV differences were found for all trunk endurance field-based tests, except for the DEE and SB-R tests. Differences

between circa- and post-PHV were only found in females for BTC and Ito tests. For all trunk endurance field-based tests, males presented higher performance than females in circa-PHV, although for BS and SB tests the results are not relevant. Sex-related differences were also found for the DEE and SB tests between boys and girls post-growth spurt. In conclusion, pre-pubertal boys and girls showed similar performance in all field-based tests, when they entered the growth spurt the differences were relevant (BTC, Ito, DEE), with boys showing higher scores than girls in all tests. After the PHV, the differences were maintained for the lateral trunk flexors (SB test) and the DEE test and were equalized for the trunk flexor tests (BTC and Ito tests) and the BS test.

CAPÍTULO I

PREÁMBULO

1. Preámbulo

La presente tesis doctoral pretende aportar información sobre diferentes aspectos relacionados con el dolor de espalda en niños y adolescentes de la Región de Murcia. Entre los diferentes aspectos estudiados se encuentra el llamado protocolo de valoración "Fitness Postural". Este protocolo de valoración evalúa los principales factores biomecánicos relacionados con el dolor de espalda como son: el morfotipo sagital integral de la columna vertebral, el rango de movimiento de la musculatura de la cadera y la resistencia de la musculatura del tronco. Con ello se pretende identificar y describir los posibles factores físicos asociados con el dolor de espalda en niños y adolescentes. Por otro lado, la presente tesis doctoral ha intentado mejorar los conocimientos acerca de las pruebas de campo que existen en la actualidad para la valoración de la resistencia muscular del tronco.

Esta información pretende servir de utilidad principalmente para profesores de Educación Física, entrenadores y todas aquellas personas que trabajen con niños y adolescentes y estén interesados en mejorar su salud postural, para poder llevar a cabo de forma adecuada propuestas y medidas de prevención óptimas para los escolares en desarrollo.

1.1. Estructura del Trabajo de Investigación

El presente trabajo de investigación se divide en dos líneas de investigación principales que guardan una estrecha relación entre sí y que a su vez se dividen en diferentes partes cada una (Figura 1).

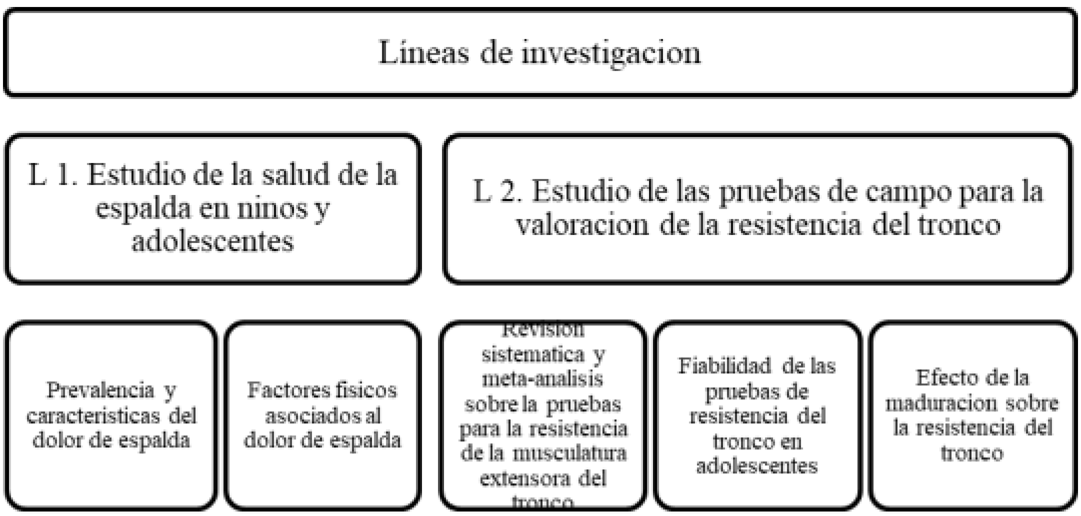


Figura 1. Partes del trabajo de investigación

El contenido del trabajo se estructura en 5 capítulos a los que se unen un apartado de referencias bibliográficas del capítulo I y otro apartado para anexos.

El primer capítulo es introductorio, y está formado por el preámbulo y una justificación de la tesis. En él se describe la estructura de la tesis y las partes que la forman, así como los requisitos para la obtención de la Mención Internacional.

En el segundo capítulo, se plantean los objetivos generales de la tesis y los específicos de cada estudio, así como las hipótesis planteadas al inicio de la investigación.

Dado que cada estudio presenta un método diferente, no se ha incluido un capítulo general sobre el método pues en el tercer capítulo se presenta toda la información de forma específica para cada estudio.

En el tercer capítulo se presentan los resultados de cada una de las líneas de investigación en formato de "estudios", los cuales están redactados en inglés y siguen la estructura de un artículo: introducción, método, resultados, discusión, conclusiones y referencias.

En el quinto capítulo se presentan las conclusiones en español e inglés, y se proponen diferentes líneas de investigación asociadas a los resultados encontrados en la presente tesis.

Finalmente, la tesis doctoral finaliza con las referencias bibliográficas de la parte introductoria y los anexos.

El compendio de estudios científicos que forman parte esta tesis doctoral, han sido seleccionados de entre todos los que se han desarrollado durante todo el período de formación, los estudios no incluidos en la presente tesis doctoral se referencian en el capítulo V.

1.2. Tesis Doctoral con Mención Internacional

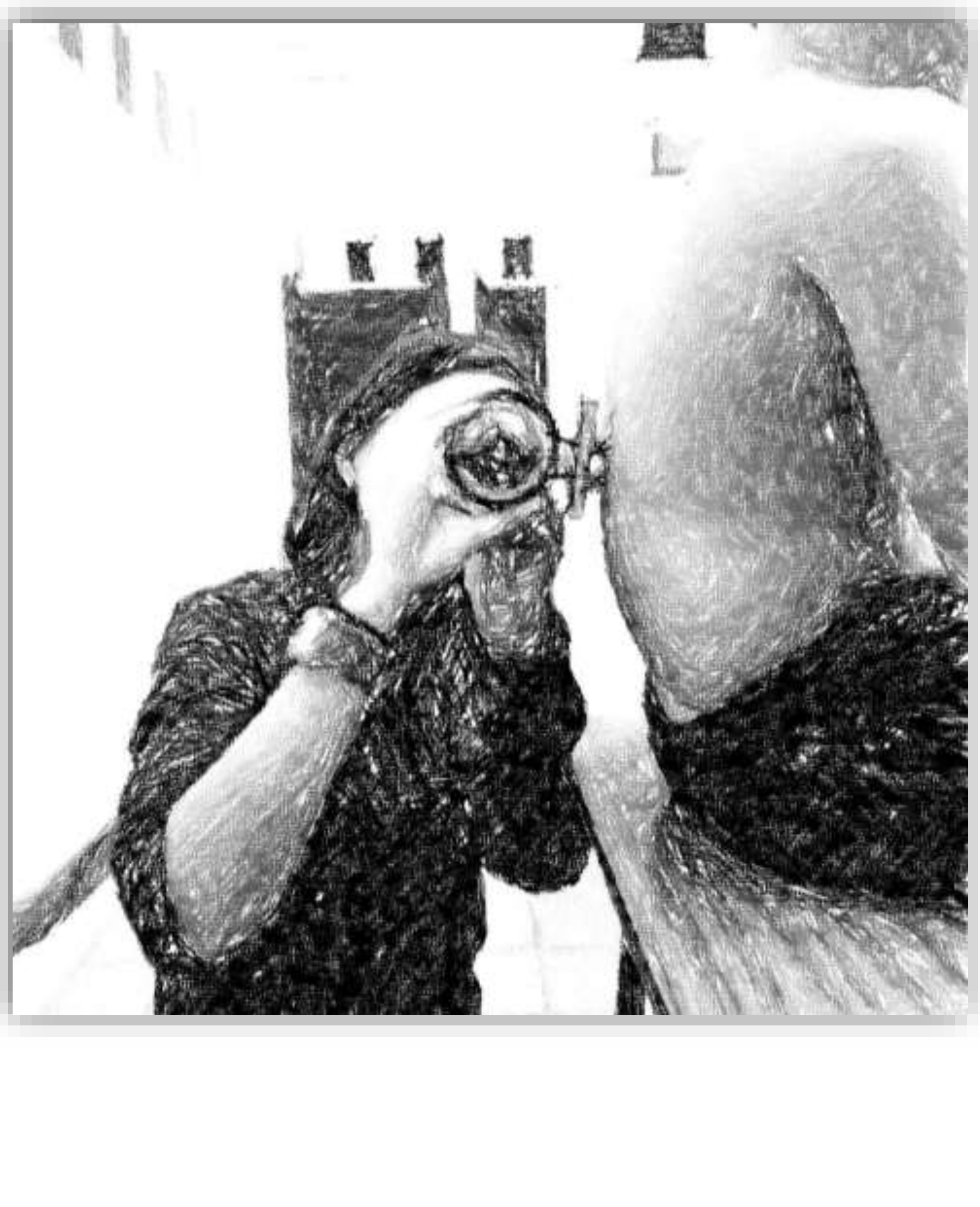
De acuerdo con la normativa establecida por la Comisión General de Doctorado de la Universidad de Murcia, en lo relativo a los requerimientos formales y de estilo que deben seguir las tesis doctorales para poder obtener la Mención Internacional, la presente tesis cumple con los siguientes criterios:

a) Como alumna de doctorado y contratada predoctoral FPU, realice una estancia breve de tres meses fuera de España en la Universidad de Gloucestershire, concretamente en la Facultad de Ciencias del Deporte y el Ejercicio desde el 1 de mayo de 2018 hasta el 1 de agosto de 2018, bajo la tutela del doctor D. de Ste Croix. En dicha estancia se llevaron a cabo los estudios 3 y 4 de la presente tesis doctoral. Asimismo, se realizaron diversos seminarios de formación organizados por la Facultad de Ciencias del Deporte y el Ejercicio de la Universidad de Gloucestershire.

b) Parte de la tesis doctoral se encuentra redactada en otro idioma diferente al español, en este caso en inglés. En concreto, aparecen redactados en inglés el resumen, todos los estudios que componen la tesis y las conclusiones.

CAPÍTULO I

INTRODUCCIÓN



1. Introducción

1.1. Dolor de espalda en niños y adolescentes

El dolor de espalda (DE) es un motivo de consulta muy frecuente entre los adultos, aunque cada vez es más habitual en niños, especialmente en niños mayores y adolescentes, como han puesto de manifiesto diversos estudios epidemiológicos que muestran prevalencias de vida que van desde el 1% a los 7 años, hasta el 12-40% a los 12 años, aumentando y casi duplicándose de los 12 a los 15 años (39-71%), coincidiendo con el inicio de la pubertad. De tal forma que al final de la adolescencia, la prevalencia del DE se aproxima a la de los adultos [1,2].

Las tasas de prevalencia varían en la literatura del 0,8 al 84%, este amplio rango de porcentajes se debe a los tipos de prevalencia analizados, las diversas definiciones y delimitaciones anatómicas de DE utilizadas, el método empleado para recopilar la información y a la diferencia en los grupos de edad estudiados [3–5].

La prevalencia indica el número de individuos que presentan una enfermedad, resultando una medida útil para conocer la magnitud de un problema. La prevalencia se puede medir como puntual (número de personas que presentan una enfermedad en un momento determinado), de período (número de personas que padecen un problema en un intervalo de tiempo específico, por ejemplo en el último mes), o de vida (número de personas que en algún momento de su vida y hasta el momento de la evaluación han padecido dicha enfermedad) [3,6].

El DE se informa de manera variable en función del artículo, encontrando términos como “dolor lumbar”, “dolor torácico”, “dolor en la zona superior”, “dolor dorsal”, “dolor de cuello” o “dolor de espalda” genérico. En otras ocasiones, se incluyen combinaciones de diferentes áreas, por ejemplo, “dolor de cuello y hombro”, “dolor de cuello y torácico”, lo que dificulta el conocimiento exacto de la prevalencia para cada zona de la columna [5,7,8]. Sin embargo, aportar una imagen o maniquí donde se delimitan las diferentes zonas de la espalda, suele ser de gran ayuda a la hora de identificar la zona del dolor, sobre todo si se trabaja con niños o adolescentes, además de incluir una definición precisa de la zona anatómica [6,9]. En una reciente revisión sistemática sobre los factores de riesgo para el dolor en el cuello y la zona torácica en jóvenes de 10 a 18 años, se encontró que las definiciones variaron considerablemente de un estudio a otro y que ninguno de los estudios proporcionaron una figura que ilustrase las áreas de interés del dolor [10] siendo esto último una limitación importante del estudio.

Por lo tanto, la claridad en la definición del DE en los estudios de prevalencia debe ser primordial para poder hacer comparaciones válidas, así como para mejorar la compresión de dicho problema de salud al aumentar el valor de los estudios individuales, además de facilitar la sinergia de la investigación internacional [11]. Por otro lado, es necesario incluir en dichas definiciones las características del dolor estudiado (frecuencia, intensidad, consecuencias, entre otros) para poder conocer mejor la aparición y el desarrollo de los primeros casos de DE, puesto que la prevalencia y severidad del DE, así como la discapacidad funcional, están relacionados con la intensidad del dolor y aumenta con la edad, de este modo se podrán llevar a cabo programas de prevención más efectivos [12–14].

Para la recopilación de la información se han empleado diferentes métodos como las entrevistas presenciales o por teléfono, los cuestionarios o a través de mensajes de texto a través del móvil. El entorno de administración (en casa o en el centro educativo) también influye, al igual que la forma en la que se presenta la prevalencia (por zona de la columna, el sexo, la edad cronológica, las etapas de maduración o el curso escolar), afectando todos estos factores en la variabilidad y validez de los estudios de prevalencia [7,8,12,15].

El DE generalmente sigue un curso recurrente a lo largo de la vida de quienes lo sufren, y el principal factor de riesgo asociado con la presencia de DE es haber tenido un episodio previo [1,2]. Con respecto al comportamiento del DE durante la infancia y la adolescencia, Aartun y colaboradores (2014) estudiaron la prevalencia e incidencia a dos años, así como la evolución, frecuencia e intensidad del dolor de espalda (cuello, zona torácica y zona lumbar) en jóvenes de 11 a 13 años, encontrando que el DE afectó a casi 9 de cada 10 participantes de 11 a 15 años. Entre los escolares que al principio del estudio no mostraron DE, alrededor de la mitad experimentaron dolor dos años después. Para la mayoría de los participantes, el dolor fue leve, poco frecuente y de baja intensidad. Sin embargo, el 14-20% refirió dolor de forma más frecuente y de mayor intensidad. El curso de dos años mostró un desarrollo progresivo del dolor e, independientemente de la ubicación inicial del mismo, éste se extendió a más zonas en los años siguientes [16]. Por otro lado, el DE es el síntoma número 3 detrás del dolor de cabeza y el dolor de estómago en pacientes jóvenes con dolor crónico (duración del dolor durante más de 3 meses). La clasificación del dolor crónico no se centra en el nivel del dolor como síntoma, sino en sus secuelas: puede no tener impacto en las actividades de la vida diaria (nivel 1), puede impedir la participación en actividades deportivas (nivel 2) o puede causar ausencias en la escuela (nivel 3) [17,18].

El DE presenta un importante impacto en los pacientes y sus familias, en las comunidades, en los sistemas de atención sanitarios y en las empresas, debido a las limitaciones en la actividad laboral, las restricciones en la participación en actividades de la vida diaria, la carga del trabajo, el uso de los recursos sanitarios y la carga financiera [6]. El impacto social y económico del DE suele variar dentro de una población en función de la situación socioeconómica, el acceso general a los servicios de salud, la distribución ocupacional y la percepción del dolor, entre otros [6].

Otro aspecto relevante relacionado con el DE, es el gasto económico derivado. En los adultos incluye los gastos de atención médica, el pago de bajas laborales, la pérdida de productividad, la rehabilitación de los empleados, los gastos administrativos y judiciales, así como de apoyo social, entre otros [6,9]. Cuando el DE se da en niños o adolescentes, el gasto económico incluye los costos sanitarios, la pérdida de horas de trabajo de los progenitores o tutores, y los anteriormente citados que aparecerán cuando sean adultos. Según el estudio de la organización "Global Burden Disease" (GBD) de 2017, la prevalencia puntual global de DE limitante de la actividad fue del 7,3%, lo que implica que 540 millones de personas se vieron afectadas en un momento dado, siendo la causa número uno de discapacidad a nivel mundial. La mayoría de las personas con DE tienen bajos niveles de discapacidad, pero el efecto aditivo de éstos, combinado con una alta discapacidad en una minoría sustancial, resulta en una carga social muy alta. En los países de ingresos altos, el DE incapacitante está relacionado con el nivel socioeconómico, la satisfacción laboral y el potencial de compensación monetaria. En Europa, el DE es la causa más habitual de baja por enfermedad y jubilación anticipada con certificación médica [19].

1.2. Localización del dolor de espalda en niños y adolescentes

El DE puede aparecer en diferentes zonas, una de las más habituales es la parte inferior de la espalda, conocida habitualmente como dolor lumbar (DL) [10]. Dionne y colaboradores (2008) a través de un proceso Delphi llegaron a un acuerdo internacional sobre una definición uniforme de DL para su uso en estudios de prevalencia: "dolor que aparece entre la parte inferior de la 12ª costilla y los pliegues glúteos que es lo suficientemente fuerte para limitar las actividades habituales o cambiar la rutina diaria por más de 1 día, puede ser irradiado o no a la pierna, y no incluye el dolor de una enfermedad febril o la menstruación" [11].

El DL puede ser agudo, de inicio brusco y de dolor intenso, generalmente desencadenado por factores físicos (algún esfuerzo, mal gesto o gesto repetitivo), por factores psicosociales (estar fatigado o cansado), por una combinación de ambas (estar distraído mientras se hace un esfuerzo) o sin causa aparente de no más de 2 semanas de duración [20–22]. En la mayoría de los casos de DL agudo no se consigue identificar un desencadenante, no se conoce una causa estructural o anatómica clara, por lo que se suele denominar DL inespecífico [22,23]. Si el dolor se mantiene más de 2 semanas y menos de 3 meses, se trata de DL subagudo y no resulta tan doloroso. Si el dolor dura más de 3 meses, se denomina crónico que puede ser continuo o con intervalos más o menos largos de remisión y de intensidad leve o moderada [20,21].

El DL en la infancia y la adolescencia se caracteriza por una gran intermitencia y tendencia a reaparecer con mayor intensidad [3]. La evolución del DL en niños y adolescentes presenta algunas similitudes con la evolución del DL en los adultos, donde la mayoría de las personas se recuperan rápidamente de nuevos episodios y la recurrencia y fluctuación es común, y sólo en una pequeña proporción el dolor se vuelve persistente e incapacitante [4,23,24]. En una reciente revisión sistemática, se evaluó el curso natural del DL en la población general desde la niñez hasta la adolescencia y encontraron 3 patrones: 1) la mayoría de los niños y adolescentes (49-53%) informaron de ninguna o poca probabilidad de experimentar DL, 2) un segundo grupo informó fluctuaciones de DL (16-37%), mientras que 3) una minoría (<1-10%) informó repetidamente de DL [4].

A pesar de que la presencia de dolor en otras zonas de la espalda ha sido menos estudiada, la evidencia sugiere que el dolor cervical, definido como el dolor en la zona posterior del cuello, y el dolor en la zona torácica, dolor en la zona media de la espalda entre T1-T12, también deben ser analizados durante la infancia y la adolescencia debido a que su incidencia y prevalencia también son elevadas [10,25,26]. En adolescentes de entre 15 y 19 años, el DL y el dolor cervical se encuentran en el top ten de los problemas de salud asociados a discapacidad en todo el mundo y están por encima de algunos problemas de salud bien reconocidos de la adolescencia, como el abuso de alcohol y drogas [27]. A los 9 años, la prevalencia de dolor en la zona torácica incluso supera a la prevalencia de dolor lumbar, igualándose a los 15 años [28,29].

Estas cifras son de particular relevancia ya que varios estudios han demostrado que la experiencia del dolor en la infancia y la adolescencia tiene repercusión sobre la experiencia de dolor más adelante en la vida [30,31]. Para el DL, por ejemplo, en un estudio longitudinal de 8 años se encontró que presentar DL a una edad temprana fue un predictor significativo de DL más adelante en la vida. Además, también se observó una asociación dosis-respuesta entre el número de días con dolor al inicio del estudio y la aparición de DL en el seguimiento, es decir, a más días con DL, mayor fue el riesgo de presentar dolor en el futuro [14]. Lo mismo parece ocurrir con el dolor torácico, en un estudio con 58 niños que informaron dolor torácico, el 90% volvieron a informar dolor después de alcanzar la madurez esquelética [32].

Dado que, no solo el DL, sino también el dolor cervical y torácico son habituales en la infancia y la adolescencia, observándose un aumento de sus prevalencias con la edad, es probable que en estos años sean más vulnerables para la aparición de un primer episodio de DE [4,7,8,10,27]. Siendo este el caso, y dado que numerosos estudios longitudinales coinciden en que la presencia de DE en la adolescencia es un predictor para padecer dicha condición en la edad adulta [3,14,24,33,34], parece ser que la aplicación de intervenciones durante la infancia o la adolescencia podría ser la única oportunidad para una verdadera prevención temprana del DE [1,2].

1.3. Factores de riesgo y/o desencadenantes de episodios de dolor de espalda en niños y adolescentes

Aunque la etiología del DE es diversa y multifactorial, numerosos estudios han intentado investigar una gran variedad de factores de riesgo (personales, psicológicos, biomecánicos y genéticos) para el DE en niños y adolescentes. La identificación de aquellos factores que pueden desencadenar o agravar el DE en jóvenes es fundamental para poder diseñar medidas preventivas eficaces e identificar qué sujetos se encuentran en situación de mayor vulnerabilidad para sufrir esta condición [10,35,36].

Previamente, es importante distinguir entre un factor de riesgo para el DE y un factor asociado con el DE. Por un lado, es necesario tener en cuenta que un factor de riesgo es cualquier rasgo, característica o exposición que está relacionado causalmente con la aparición de un cambio en una condición de salud relevante, es decir, el factor de riesgo debe estar presente, como mínimo, antes de la aparición de la enfermedad para que se considere causal. Si se mide un factor de riesgo al mismo tiempo que una enfermedad, se infiere asociación y no causalidad, salvo que se sepa con certeza que el factor de riesgo existía antes del inicio de la misma. Por lo tanto, otros factores estudiados pueden no haber ocurrido antes del inicio de la enfermedad y pueden tener una relación bidireccional con ella [8,37,38].

Con respecto al DE, el primer caso de dolor sería el inicio de la enfermedad, y un episodio de DE sería un evento nuevo de DE una vez que la enfermedad ya ha ocurrido. Por lo tanto, un factor de riesgo sería aquel que causa la enfermedad (primer caso de DE), mientras que un desencadenante daría lugar a un nuevo episodio de DE [8]. Por tanto, también es razonable postular que el abordaje preventivo puede ser diferente para prevenir el inicio de la enfermedad (a partir de los factores de riesgo), que para prevenir nuevos episodios de la misma (a partir de los factores desencadenantes) [38].

Otro aspecto a considerar cuando se estudian los factores de riesgo o desencadenantes es que las asociaciones individuales pueden ser relativamente débiles, pero es posible que la combinación de factores o la suma de factores aumente el riesgo de DE debido a que podrían interactuar multidireccionalmente entre sí. Por lo tanto, el estudio de los factores de riesgo debería realizarse mediante un enfoque basado en relaciones de dependencia e independencia condicional entre variables dentro de un contexto multivariante y dinámico, y permitir así el desarrollo de programas preventivos mucho más eficaces [37,39]. En esta línea, existen ciertos factores de riesgo que podrían predisponer a los jóvenes a sufrir DE (factores de riesgo intrínsecos o no modificables) y luego, existen otros factores desencadenantes del DE (por ejemplo, posiciones de trabajo incómodas, inactividad, movimientos bruscos), la combinación de ambos podría dar lugar al desarrollo de la enfermedad. Por ejemplo, las niñas (factor 1) con un estado puberal avanzado (factor 2) podrían ser susceptibles al DE que posteriormente aparece debido a la inactividad física (factor 3). Por lo tanto, si se pueden evitar los factores desencadenantes, sería posible reducir el número de episodios de DE entre los jóvenes.

Teniendo en cuenta la literatura existente incluyendo las últimas revisiones sistemáticas y meta-análisis realizados al respecto, los factores de riesgo o desencadenantes más probables del DE en los jóvenes son predominantemente biológicos y no modificables (por ejemplo, sexo, edad, estado puberal, historia familiar, altura), lo que los convierte en objetivos no elegibles para intervenciones preventivas [37]. Por otro lado, se plantean los factores modificables y bidireccionales (por ejemplo, peso, IMC, resistencia y flexibilidad muscular, postura, actividad física, sedentarismo, calidad y cantidad del sueño, tabaquismo, entre otros), los cuales deben ser el centro de atención para la aplicación de futuras medidas preventivas. A continuación, se exponen algunos de los factores de riesgo o desencadenantes del DE en niños y adolescentes más estudiados, y algunos métodos que se emplean para su valoración (Tabla 1).

Tabla 1. Factores de riesgo o desencadenantes del DE en niños y adolescentes

Factores de riesgo no modificables	Factores de riesgo modificables
Personales	*Psicosociales*
Sexo	*Estilo de vida*
Edad cronológica	Actividad física y sedentarismo
Altura o crecimiento acelerado	Tabaquismo
Estado madurativo	Enfermedades sistémicas
Historia previa de dolor de espalda	Sueño
Nivel socioeconómico	*Biomecánicos*
Genéticos	Peso e índice de masa corporal
Historia familiar de dolor de espalda	Postura
	Flexibilidad/Movilidad
	Fuerza/Resistencia muscular
	Mochilas

1.3.1. Factores no modificables

1.3.1.1. Personales

Sexo

En general, la mayoría de los estudios de prevalencia y factores de riesgo para el DE informaron asociaciones positivas entre el sexo femenino y el dolor [2,25]. De hecho, en una reciente revisión sistemática sobre los potenciales factores de riesgo y desencadenantes del DE en niños y adolescentes, encontraron en 32 estudios de los 53 incluidos en la revisión, una asociación positiva entre el DE y el sexo femenino, en 3 estudios se encontró una mayor prevalencia de DE entre los hombres y en los 18 restantes no encontraron ninguna asociación [37]. Existe la hipótesis de que estas diferencias entre hombres y mujeres se podrían deber a la modulación del dolor debido a que durante la adolescencia los niveles de estrógenos aumentan notablemente en las chicas y altera la actividad de la serotonina, un neurotransmisor cerebral que puede influir negativamente en la percepción del dolor [40].

Edad cronológica

Al igual que ocurre con el sexo, la edad cronológica también es un predictor claro de DE. A mayor edad en los niños o adolescentes, mayor es la probabilidad de sufrir DE [2,25,36].

Crecimiento acelerado

La medición del crecimiento implica medir el tamaño del cuerpo o el tamaño de partes específicas del cuerpo, y rastrear cómo aumenta el tamaño con el tiempo. Los métodos utilizados para medir el tamaño y las proporciones corporales se denominan colectivamente antropometría y son extremadamente fiables si los emplean evaluadores expertos. El tamaño corporal total se mide de manera simple en términos de altura y masa corporal. Otras medidas incluyen longitud, anchura, circunferencia, proporción (o relación) entre dos medidas y composición corporal. La mayoría de las dimensiones corporales siguen un perfil de desarrollo similar al de la altura y la masa corporal: crecimiento rápido en la infancia y la niñez temprana, crecimiento más lento en la niñez media, crecimiento acelerado durante la adolescencia y finalmente una desaceleración seguida de la terminación cuando se alcanza el tamaño adulto. Sin embargo, el momento exacto de los brotes de crecimiento puede diferir entre los segmentos corporales [40].

Dado que las extremidades experimentan un crecimiento acelerado más temprano que la columna y el tórax, las proporciones corporales cambian continuamente al igual que el peso corporal, la fuerza muscular y la longitud de los músculos. Por así decirlo, hay un cambio continuo de la situación biomecánica individual, particularmente durante el brote de crecimiento puberal [17]. Por lo que la aceleración del crecimiento podría considerarse un período particularmente vulnerable debido a los cambios repentinos de carga mecánica en la columna [7].

El aumento de la altura como tal, presenta estimaciones inconsistentes de asociación con el DE [8]. Sin embargo, se ha visto que la aceleración del crecimiento o brote de crecimiento, con una desviación estándar de unos 4,3 cm entre los 11 y los 14 años se asocia positivamente con el DL [41]. Recientemente, se ha observado que aquellos escolares que experimentan un mayor crecimiento lineal también muestran una mayor frecuencia y duración de DE [42].

Estado madurativo

Los cambios más evidentes que ocurren durante el desarrollo puberal son físicos (aumento de la altura y cambios en la composición corporal) y hormonales (mayor secreción de estrógenos en las chicas y más testosterona en los hombres), los cuales derivan en una mayor altura y masa muscular en los chicos y un aumento de la grasa corporal en las chicas [43–45]. Por ello, se ha propuesto que este período de cambios puede afectar a la aparición del DE [7]. Las niñas comienzan la pubertad antes que los niños, lo que podría explicar por qué refieren DE antes [46].

Lardon y colaboradores (2014) encontraron en su revisión sistemática una clara asociación entre la pubertad y el DE, y la existencia de una posible relación causal. Sin embargo, no aportaron ninguna información sobre qué aspectos de la pubertad, específicamente, pueden contribuir al DE. Por lo tanto, se puede decir que existe una asociación entre el DE y aquellos que presentan un estado puberal avanzado [7,37,42,47–49].

El estado de madurez a menudo se cuantifica en términos de edad biológica, lo que proporciona una medida de cuán desarrollado está un sistema biológico en el continuo desde inmaduro hasta completamente maduro. Sin embargo, hay procesos biológicos específicos que suelen madurar a diferentes ritmos, lo que dificulta definir y medir la madurez. Normalmente, la madurez y la edad biológica se consideran en tres contextos posibles (Tabla 2): la edad esquelética, que proporciona una clasificación de la edad biológica basada en el proceso del esqueleto desde el cartílago hasta el hueso, la edad sexual, que proporciona una clasificación de la edad biológica basada en características sexuales secundarias, y la edad somática, que proporciona una clasificación de la edad biológica basada en un progreso y una tasa de desarrollo hacia el tamaño y la proporción de adultos completos [40,50].

Existe cierto debate sobre cuánto se relacionan entre sí las diferentes medidas de la edad biológica. Algunos investigadores sugieren una relación razonable, mientras que otros sugieren una relación relativamente pobre [50]. Por tanto, es necesario tener cuidado al comparar los marcadores de maduración.

Tabla 2. Ventajas, inconvenientes y métodos para la cuantificación del estado madurativo

Métodos	Ventajas	Inconvenientes
Maduración esquelética		
Se evalúa habitualmente por medio de tres métodos que se basan en la radiografía de la mano y la muñeca, pero utilizan diferentes criterios de evaluación: -Método de Greulich-Pyle -Método de Tanner-Whitehouse -Método de Fels	-Se considera el "Gold-Standard"	-Equipos especializados y costosos -Personal cualificado -Exposición a radiación -Desacuerdo entre los diferentes métodos para el cálculo de la edad esquelética
Maduración sexual		
Se evalúa observando y calificando las características sexuales secundarias (vello púbico y desarrollo de los senos y genitales). Diferentes formas: -Escala de Tanner -Edad de la primera regla -Escala de Desarrollo Puberal	-Edad de la primera regla y Escala de Desarrollo Puberal: no invasivo	-El método Tanner tiene una aplicabilidad limitada fuera de entornos clínicos debido a la invasión de la privacidad personal. -Tanto la escala de Tanner como la Escala de Desarrollo Puberal, pueden ser auto-administrados, pero los chicos tienden a sobreestimar y las chicas a subestimar su desarrollo sexual. -La edad de la primera regla está influenciada por la memoria y el sesgo de recuerdo.
Maduración somática		
En este enfoque, se toman medidas antropométricas para estimar el estado de madurez de forma indirecta. Se pueden usar diferentes modelos matemáticos: -Ecuación de Mirlwald y colaboradores (2002) [51]: aporta una medida, conocida como compensación de la madurez	-No invasivo -Fácil aplicación -Válidas cuando los sujetos están cerca del momento del PVC real.	-El nivel de error asociado con estas mediciones y la sensibilidad con la que pueden detectar la juventud de maduración temprana y tardía.

Tabla 2. Ventajas, inconvenientes y métodos para la cuantificación del estado madurativo

Métodos	Ventajas	Inconvenientes
("Maturity Offset"), que estima los años por los que el individuo está antes o después del pico de velocidad de crecimiento (PVC). -Ecuación de Moore y colaboradores (2015) [52]: aporta la misma medida, pero a partir de la altura y la edad del sujeto. -Índice de madurez de Fransen y colaboradores (2018) [53].		-Las ecuaciones parecen sobrestimar la edad en el PVC en chicas.

Historia previa del dolor de espalda

Como se ha comentado previamente, es bien sabido que haber sufrido DE durante la adolescencia es un fuerte factor de riesgo de DE en la edad adulta [37]. En el estudio de Hestbaek y colaboradores (2006), encontraron correlaciones entre el DL en la adolescencia y el DL en la edad adulta, especialmente para el DL persistente, que demostró un aumento de riesgo cuatro veces mayor [14]. Del mismo modo, en un estudio con adolescentes de 14-16 años se encontró que los informes de DL durante esas edades son fuertes predictores de DL futuro 3 años después [54]. Además, en otros estudios se vio que tener dolor cervical a los 14 años, aumentaba las probabilidades de DL, destacando que para muchas personas es coexistente con otras áreas de dolor [55–57].

Nivel socioeconómico

Existen resultados variados en cuanto a esta variable. En una reciente revisión sistemática encontraron 15 estudios que analizaban dicho factor, de los cuales 7 estudios informaron asociaciones entre ciertos factores socioeconómicos (menor índice socioeconómico, menor clase social, menor nivel de educación de los progenitores, zonas urbanas, entre otros) y un mayor riesgo de DE, mientras que 8 estudios no informaron de ninguna asociación [37].

Historia familiar del dolor de espalda

Algunos estudios longitudinales realizados en niños y adolescentes, han mostrado que existe una evidencia de calidad moderada de que los niños y adolescentes con antecedentes familiares de DE tienen un 58% más de probabilidades de experimentar DE que los niños de familias sin antecedentes de dolor [58]. En los análisis por subgrupos de esta misma revisión sistemática, vieron que los niños con antecedentes de dolor materno tenían un 53% de probabilidad de dolor, frente a un 59% si procedía del padre y un 99% si procedía de un hermano.

1.3.1.2. Genéticos

La manifestación de DE se ha visto influenciada por factores genéticos, estimándose la heredabilidad de DL superior al 40% [59]. En este estudio de gemelos, se vio que a medida que los adolescentes se vuelven adultos, la influencia del entorno no compartido aumenta en relación con los componentes genéticos. Un componente importante de este efecto podría estar relacionado con el trabajo cuando los gemelos adultos obtienen un empleo diferente [59].

1.3.2. Factores modificables

1.3.2.1. Psicosociales

Los factores psicosociales "positivos" generalmente se han encontrado asociados con el DE, mientras que los factores "negativos" han sido relacionados con el aumento del DE [8,10,60–62], por ejemplo depresión, estrés, baja calidad de vida, pobre salud mental, mala valoración de su condición física o baja autoestima.

En un estudio realizado con adolescente daneses [60], se vio que aquellos que indicaban frecuencias más altas en los diferentes factores psicológicos (¿con qué frecuencia se ha sentido deprimido/de mal humor/nervioso/con dificultad para dormir?) tenían más probabilidades de informar DE. Por otro lado, aquellos que informaron DE también tenían niveles significativamente más altos de soledad y más bajos de aceptación de los compañeros en comparación con los que no tenían DE.

Stallknecht y colaboradores (2017), encontraron que los adolescentes daneses de entre 10 y 14 años con valores medios y altos de estrés tenían una mayor probabilidad de informar DE en comparación con los que no informaron estrés [61]. Los adolescentes que informaron un peor bienestar general también tuvieron mayores probabilidades de informar DE en comparación con aquellos que informaron un mejor bienestar.

Entre los cuestionarios más utilizados para la valoración de los factores psicosociales se encuentran: "Strengths and Difficulties Questionnaire", "SF-36 health questionnaire", "Youth Self Report", "Beck Depression Inventory for Youth", "Perceived Self-Efficacy", "Mental Health Inventory", "Kidscreen".

1.3.2.2. Estilo de vida

Actividad Física y sedentarismo

Es bien sabido que la actividad física tiene un efecto positivo en la salud al disminuir la adiposidad, mejorar la salud y capacidad cardiovascular, la salud mental, el rendimiento académico, la salud músculo-esquelética, la condición física y la densidad mineral ósea [63,64]. Sin embargo, en la literatura científica se ha observado que la asociación de la actividad física con el DE genera resultados con una curva en forma de "U" [65–67].

Por un lado, la falta de actividad física y un estilo de vida sedentario están considerados como factores de riesgo para el dolor de espalda. En este sentido, Szita y colaboradores (2018) observaron que los adolescentes que indicaban DE pasaban más tiempo viendo la televisión que aquellos sin DE [68]. Martínez-Crespo y colaboradores (2009) también encontraron asociación entre el DE y las horas que los adolescentes pasaban sentados con el ordenador o la televisión (el 45% pasaba más de 2 horas) [69]. Cabe añadir que existe una tendencia hacia el sedentarismo conforme aumenta la edad de los adolescentes, Szita y colaboradores (2018) encontraron que los jóvenes mayores de 12 años, pasaban mucho más tiempo sentados en comparación con la población más joven, siendo posible que el uso creciente de aparatos electrónicos les lleve a pasar más tiempo sentados en comparación con otras generaciones anteriores [68,69].

Y por otro lado, la participación intensa en actividades físicas, más de 6 horas a la semana, se asoció con DL en jóvenes de ambos sexos [70]. El impacto biomecánico de una actividad deportiva específica, la cantidad de horas de entrenamiento y la resistencia mecánica relacionada con la edad es más relevante que el sexo. En este sentido, en un gran estudio epidemiológico (26.766 participantes) el entrenamiento semanal de más de 6 horas condujo a más DE que cualquier actividad o deporte moderado [17,71]. El DE en un deportista puede ser independiente del deporte o causado total o parcialmente por él. Por ejemplo, deportes donde existen movimientos de rotación, de hiperextensión lumbar (gimnasia, golf, rugby, bádminton o voleibol), son los que tienen el mayor riesgo de sufrir DE [71]. Las cargas elevadas que actúan sobre las placas de crecimiento de las vértebras relativamente débiles y susceptibles, dan como resultado cambios en el perfil sagital de la columna vertebral. Por ello, los controles médicos deportivos deberían ser obligatorios en los deportistas activos [17,72–74].

Por lo tanto, una mala condición física o los deportes de alto rendimiento, parecen ser factores de riesgo para el desarrollo del DE más prevalente, particularmente durante el estirón puberal. Las chicas y determinadas actividades deportivas que incluyen momentos repetitivos de hiperextensión-rotación lumbar o cargas axiales elevadas tienen un mayor riesgo [17], lo que sugiere que los niveles moderados de actividad física pueden ser más recomendables para prevenir el DE. En el caso de jóvenes muy activos, habría que centrarse en la intensidad, la recuperación, la planificación del entrenamiento, entre otras cosas, para reducir el riesgo de DE [70].

Cabe destacar que en algunos estudios se encuentran resultados contradictorios [37,75], principalmente por la dificultad para cuantificar la actividad física (auto-informada mediante cuestionarios como el "IPAQ" o de forma objetiva a través de acelerómetros) en niños y adolescentes. Por ejemplo, los niños inactivos tienden a sobrestimar su tiempo en el auto-informe de actividad física; por otro lado, los niños tienden a moverse en períodos cortos de actividad física intensa que dura menos de 15 segundos, lo cual es un desafío el poder capturar tal actividad [76]. Además, es importante clasificar la actividad física como un continuo de niveles y no de forma dicotómica, activo o inactivo [67].

Tabaquismo

En una reciente revisión sistemática, donde se recopilaron los estudios que analizaban la asociación entre el DE y el hábito de fumar en adolescentes, observaron que en todos los estudios incluidos encontraban asociaciones positivas entre el DE y fumar [37]. En uno de los estudios longitudinales se vio que fumar a los 14 años fue un predictor de DE a los 17 años [77]. Algunos autores atribuyen estas asociaciones a la influencia de diferentes características psicológicas y sociales negativas, junto con otros factores como sedentarismo, malos hábitos alimenticios y obesidad, entre otros [77–80].

Enfermedades sistémicas

Se ha visto que tener una enfermedad crónica presenta una fuerte asociación con el DE, algunas enfermedades fueron asma, migrañas, dolor abdominal y resfriados/enfermedades leves. Estas pueden ser comorbilidades del DE, lo que significa que una podría ser precursora de la otra o podrían tener una causa común [8].

Hábitos de sueño

Con respecto a los hábitos de sueño, se encuentran asociaciones positivas entre el DE y el sueño insuficiente [37,60,68]. En el estudio de Auvinen y colaboradores (2010), encontraron que la cantidad y calidad insuficiente de sueño a los 15-16 años predijo el dolor cervical y lumbar en las chicas de 18-19 años [81]. Otros autores encuentran que la incomodidad y los trastornos del sueño, son un predictor de dolor en más de una zona de la espalda y también del dolor frecuente [13].

Algunas explicaciones podrían ser los cambios en las rutinas diarias de muchos adolescentes por la exposición a las pantallas en las horas previas a dormir, hacer deberes y trabajos por la noche, el consumo de cafeína, nicotina o de otras drogas.

1.3.2.3. Biomecánicos

Peso e índice de masa corporal

La literatura existente informa de asociaciones inconsistentes entre el DE durante la infancia y la adolescencia y el índice de masa corporal (IMC) [41,80,82–84]. Una revisión sistemática que incluyó a 5.567 participantes categorizados en peso normal o sobrepeso, concluyó que los niños con sobrepeso tienen un mayor riesgo de DL con una razón de riesgo de 1,42 (IC del 95%: 1,03-1,97), sin embargo, los estudios incluidos ofrecían solo evidencia de baja calidad [85]. En cambio, en la revisión de Beynon y colaboradores (2019), de los 8 estudios que analizaron el IMC y el DE, 3 encontraron una mayor prevalencia de DE conforme aumentaba el IMC y 5 no encontraron asociación, por lo que concluyeron que no hay evidencia suficiente para concluir que existe una relación entre el IMC y el DE [8].

A pesar de las diferencias en los estudios, Palmer y colaboradores (2020) en un estudio con casi medio millón de niños, encontraron que los niños con sobrepeso u obesidad tenían un mayor riesgo de desarrollar DE en la columna lumbar, torácica y cervical, después de ajustar por factores de confusión. Cabe añadir que se trata de dolor severo, pues los niños del estudio acudieron a su centro médico y el pediatra confirmó el diagnóstico [86]. Se desconocen las razones del aumento del riesgo de DE en niños con sobrepeso u obesidad, pero la etiología puede ser mecánica, biológica o psicológica [23]. Al no encontrar asociación entre el peso al nacer y el DE, los autores sugirieron que el aumento de peso en una edad temprana, aumentaba el riesgo de DE [86].

Postura

Desde un punto de vista preliminar, se ha visto que la presencia de desalineaciones en la columna (por exceso o por defecto) y ciertas posiciones al sentarse (mala higiene postural) se asociaron con DE [10,35,37].

La aparición de desalineaciones en la columna vertebral en bipedestación se asoció con el DL, al igual que el aumento del ángulo lumbo-pélvico [87–89]. En el estudio prospectivo de Smith y colaboradores (2008), cualquier desviación en la columna vertebral (hipercifosis, dorso-plano o hiperlordosis) era un riesgo para el DE [89]. Los estudios de Dolphens y colaboradores (2012, 2013), informaron que la postura hipercifótica se asociaba con el DL en niños [87,88]. En otro estudio reciente del mismo grupo [90], vieron que tanto una mayor lordosis lumbar como un aumento en la retroversión de la pelvis en bipedestación fueron dos factores asociados al DL.

Las desalineaciones posturales pueden proceder de los cambios que se producen durante el crecimiento, especialmente durante el estirón puberal, que puede resultar en un crecimiento desigual de las vértebras (acuñamientos) o un desarrollo desequilibrado de la musculatura del tronco [17,91]. Es importante destacar que la mayoría de los malos hábitos posturales se establecen durante las fases de crecimiento y desarrollo madurativo, provocando cambios posturales, en particular cambios en el plano sagital y frontal, que también pueden estar relacionados con la postura sentada, ya sea al asistir a la escuela o al mirar la televisión [17,91].

Por otro lado, se vio que las desalineaciones sagitales de la columna en bipedestación se asociaron con un mayor grado de hipercifosis en sedentación [92], y que esa mayor hipercifosis sentado se asociaba débilmente con una mayor tasa de DE. Se ha sugerido que una postura tóraco-lumbar flexionada en sedentación puede aumentar la carga espinal, lo que representa un mecanismo potencial para la aparición del DL [93].

Cabe decir que también se han encontrado resultados contradictorios debido a los diferentes métodos que se emplean para la valoración de la columna vertebral y las marcas de referencia que se utilizan, los cuales se pueden ver resumidos en la Tabla 3.

Tabla 3. Técnicas para la valoración de la disposición sagital de la columna vertebral

Técnica/Método	Procedimiento
Mediciones angulares	
Radiografía (Gold-standard) [94]	El método de Cobb es el más difundido para la medición de las curvaturas de la columna, tanto en los planos frontal y sagital. Se mide localizando las vértebras que constituyen el límite superior e inferior de las curvas, las que más se inclinan hacia la concavidad de la misma (las más inclinadas con respecto a la horizontal). Una vez identificadas, se traza unas líneas que prolongan el platillo superior de la vértebra del límite superior y el platillo inferior de la vértebra del límite inferior. El valor angular se obtiene de la intersección de las dos líneas. En el caso de que no haya espacio para la intersección, el valor angular se obtiene trazando las líneas perpendiculares de éstas.

Tabla 3. Técnicas para la valoración de la disposición sagital de la columna vertebral

Técnica/Método	Procedimiento
Fotografías con marcadores retro-reflectantes [89]	Se colocan marcadores retro-reflectantes en 7 puntos: ángulo lateral del ojo, trago de la oreja, apófisis espinosas de C7 y T12, espina ilíaca antero-superior y trocánter mayor. A partir de dichos puntos se sacan las diferentes medidas angulares con un software de análisis de imágenes.
Inclinómetro [95,96]	Previo a la valoración, se marca en la piel la apófisis espinosa de la primera vértebra dorsal (T1), la transición tóraco-lumbar (T12-L1) y quinta vértebra lumbar (L5). En bipedestación, para la curva dorsal, el inclinómetro se coloca al inicio de la curvatura torácica, calibrándose a 0°, y se contornea la zona hasta alcanzar el mayor valor angular. Para la curva lumbar, se calibra a 0º en el punto donde se determina la mayor curvatura dorsal y se vuelve a contornear hasta alcanzar la mayor angulación para la zona lumbar. La valoración del morfotipo sagital integral incluye las posiciones de sedentación asténica y flexión del tronco. Para la valoración de la columna en esas posiciones, se coloca el inclinómetro en las marcas realizadas previamente.
Pantógrafo espinal [95,97]	Es un dispositivo articulado que sirve para trazar a escala las curvas cifótica y lordótica. El examinador desliza la rueda del pantógrafo friccionando levemente los procesos espinales entre C7 y L5. Los ángulos de cifosis y lordosis se miden después trazando unas líneas tangentes al contorno de las curvas, las líneas elegidas son las que presentan la máxima desviación con la vertical, los ángulos se obtienen con sus intersecciones. De este modo, dibujando tres líneas se pueden obtener los grados de cifosis torácica y lordosis lumbar.
Cifómetro de Drebunner [95,98,99]	Este instrumento tiene un transportador de ángulos al final de los dos brazos paralelos dobles conectados a dos bloques que realiza un análisis del contorno de la superficie de la espalda. Este instrumento puede medir ángulos de hasta 70° con un margen de error por debajo de ±1°. La cifosis (marcada con un signo positivo) se mide desde un punto entre las apófisis espinosas de las vértebras T2 y T3 y desde un segundo punto entre T11 y T12. La lordosis (marcada con un signo menos) se mide de manera similar entre T11 y T12 y S1 y S2.
Regla flexible [95,100]	Consiste en colocar una regla flexible sobre el perfil sagital de la columna y moldearla a la forma de éste. Para la medición de los ángulos, se marcan primero en la piel del sujeto los ápex de las curvas raquídeas. Se aplica la regla sobre la superficie de la columna y se moldea aquella a la forma del perfil sagital. Se dibuja sobre un papel el contorno de la regla moldeada obtenido, marcando también la señal central y las extremas. Se trazan luego

Tabla 3. Técnicas para la valoración de la disposición sagital de la columna vertebral

Técnica/Método	Procedimiento
	dos líneas rectas que unan los extremos de cada curva y se mide el ángulo formado entre ellas.
Raster-Estereografía [101]	Es un método de escaneo óptico de luz para medir superficies, el sistema consiste en un proyector de luz que proyecta una rejilla de líneas en la parte posterior del sujeto. Un software informático realiza una reconstrucción de la superficie del dorso transformando las rayas y su correspondiente curvatura en un diagrama de dispersión. A continuación, se puede calcular un modelo 3D de la columna en función de la forma convexa específica de la apófisis espinosa de la prominencia vertebral y la concavidad de los hoyuelos de Venus.
Spinal Mouse [102]	Se trata de un dispositivo electromecánico de mano asistido por ordenador que aloja acelerómetros y registra la distancia y los cambios de inclinación con respecto a la plomada a medida que se desliza por la columna vertebral. Luego, esta información se utiliza para calcular las posiciones relativas del sacro y los cuerpos vertebrales de la columna vertebral ósea subyacente utilizando un algoritmo recursivo inteligente.
Mediciones lineales	
Flechas sagitales [95]	Con una plomada y una regla se aproxima la plomada al cuerpo del sujeto y se miden 4 distancias: cervical (C-7), torácica (zona más convexa de la espalda), lumbar (parte más cóncava) y sacra (inicio del pliegue interglúteo). A partir de estas distancias se calcula el índice cifótico y lordótico.

Flexibilidad/Movilidad

En función del estudio se encuentran diversos métodos e instrumentos para la valoración del rango de movimiento (ROM) de diferentes partes del cuerpo que podrían estar asociadas con el DE, como por ejemplo de la musculatura flexora o extensora de la cadera o la movilidad global de la columna vertebral [103,104]. Es por ello que se suelen encontrar resultados contradictorios [35].

La situación en términos de movilidad de la cadera y flexibilidad de la musculatura flexora (psoas-ilíaco o cuádriceps) o extensora (musculatura isquiosural) se vuelve compleja ya que los métodos de medición difieren sustancialmente [35]. Kanchanomai y colaboradores (2015), por ejemplo, valoraron la extensión de cadera mediante la prueba de Thomas modificada (ROM de flexores de cadera) y la extensión de rodilla a través de la prueba de extensión activa de la rodilla (ángulo poplíteo) (ROM de flexores de rodilla), encontrando asociaciones positivas entre la disminución de la extensibilidad de la musculatura flexora de la cadera y el DL [105]. Por otro lado, Jones y colaboradores (2005) también encontraron asociaciones positivas entre el DL, pero en este caso con el ROM de la musculatura flexora de la cadera medido a través de la prueba de flexión de cadera con la rodilla extendida [106].

En cuanto al movimiento de la columna en el plano sagital, se suele utilizar la flexión y extensión de la columna en los exámenes clínicos para evaluar su movilidad [107]. En un estudio prospectivo se encontró que la flexión del tronco podría estar asociada con el DL en los adolescentes [108]. En el estudio prospectivo de Aartun y colaboradores (2016), emplearon dos métodos diferentes para la valoración del ROM lumbar (la prueba de Schober y la prueba dedos-suelo), pero no encontraron asociación con el DL en adolescentes. Otros estudios también incluyen la valoración de la extensión y la inclinación lateral del tronco [106,108,109].

Sin embargo, de nuevo, vuelven a aparecer resultados en forma de "U", pues tanto una la elevada movilidad como la reducción de la misma, pueden asociarse con el DE. Por un lado, la combinación de una elevada movilidad sagital lumbar junto con una baja resistencia en la musculatura extensora del tronco, emergió como un factor de riesgo de DL en un estudio prospectivo de alta calidad [110]. Por otro lado, una reducción en el ROM lumbar, indica rigidez en la zona lumbar y puede dar lugar al desarrollo de DL (Jones et al., 2005). Se puede decir que los resultados con respecto a la movilidad espinal son bastante diversos, en parte por la variedad de métodos de medición que se emplean (prueba de Schober, distancia entre dedos-suelo o dedos-planta, flexión lumbar, extensión lumbar o inclinación lateral, entre otras).

Fuerza/Resistencia muscular

Se ha demostrado que la presencia de déficits en los músculos del tronco aumenta la carga en la columna vertebral y reduce la estabilidad espinal generando patrones alterados de reclutamiento de los músculos del tronco, que es un sello distintivo del DL [111–113]. La revisión sistemática realizada por Potthoff y colaboradores (2018), reveló que existe evidencia concluyente de que algunos parámetros físicos cuantificables, podrían estar asociados con el DE en los adolescentes [35]. Esto explica particularmente que los adolescentes con DE suelen presentar una baja resistencia en la musculatura extensora del tronco en comparación con compañeros asintomáticos [35,63,114].

En esta misma línea, varios estudios prospectivos [57,110] revelaron una mayor probabilidad de sufrir DL en los años siguientes a la valoración en aquellos jóvenes que mostraron una baja resistencia en los músculos extensores del tronco. De igual forma, la presencia de desequilibrios en la resistencia muscular del tronco (mayor resistencia en la musculatura flexora del tronco en comparación con la musculatura extensora, por ejemplo) también se asoció al DL en niños y adolescentes [106,110,115].

Por otro lado, la reducción de la resistencia de los músculos flexores del tronco, con la excepción de un estudio de baja calidad, también se asoció con el DL en tres estudios transversales, dos de los cuales se centraron en adolescentes con DL recurrente o continuo [106,116,117]. Sin embargo, en otros estudios se ha visto que una mayor resistencia en la musculatura flexora del tronco también se asociaba con el DE en adolescentes [116]. De nuevo, se ven resultados en forma de "U", los cuales pueden estar influenciados por la multitud de pruebas diferentes que existen para la valoración de la resistencia de la musculatura del tronco (Tabla 4).

Tabla 4. Recopilación de las diferentes pruebas de campo que existen para la valoración de la resistencia muscular del tronco

Musculatura	Pruebas estáticas	Pruebas dinámicas
Extensores del tronco	Prueba "Biering-Sorensen" [118] Prueba "Prone isometric chest raise" [119] Prueba "Prone Double Straight-Leg Raise" [120] Prueba "Supine Bridge" [121]	Prueba "Dynamic Extensor Endurance" [122]
Flexores del tronco	Prueba "Ito" [119] Prueba "Flexor Endurance" [112] Prueba "Prone Bridge" [123]	Prueba "Curl-Up" [124] Prueba "Partial Curl-Up" [125] Prueba "Bench-Trunk Curl Up" (BTC) [126] Prueba "Flexion-Rotation Trunk" [127]
Flexores-laterales del tronco	Prueba "Side Bridge" [112]	

Con respecto a la fuerza de la musculatura del tronco, los estudios son menos claros. Sólo un estudio transversal encontró alguna asociación en los chicos entre el DL y el ratio flexión/extensión de la musculatura del tronco a velocidades intermedias [128]. Por lo que es probable, que la fuerza muscular isocinética del tronco desempeñe un papel menos importante en la aparición del DE en niños y adolescentes [63]. De hecho, McGill propone trabajar primero la resistencia y estabilidad de la musculatura del tronco antes que la fuerza en personas con DE [111,129].

En cuanto a las extremidades inferiores, se propone que la presencia de alteraciones en éstas reduce la capacidad de absorción, aumentando las cargas sobre la columna vertebral [107,130]. Las alteraciones en el tobillo y en el pie, pueden considerarse una causa potencial de DL debido a la interrupción de la cadena cinética desde el pie hasta la espalda. Por ejemplo, una pronación excesiva del pie puede llevar a una rotación tibial y femoral interna, lo que puede favorecer una inclinación pélvica anterior [107,130]. Se han encontrado estudios donde aparecen asociaciones entre un peor rendimiento de las extremidades inferiores y el DL. Por ejemplo, los adolescentes que presentaron DL también tuvieron una menor resistencia en cuádriceps y glúteos [131,132].

Las conclusiones a las que llegaron Potthoff y colaboradores tras la revisión sistemática que realizaron, fueron que se necesitan más estudios prospectivos de alta calidad que evalúen la posible causalidad entre los factores de riesgo físicos y el DE en niños y adolescentes y recomiendan que dichos estudios se centren en la resistencia de la musculatura del tronco, preferiblemente en combinación con la movilidad de la columna vertebral sagital, la alineación postural sagital y la neurodinámica como posibles factores de riesgo modificables para el desarrollo de DE en la infancia y adolescencia [35].

1.3.2.4. Mochilas

Los resultados de la literatura sobre la asociación entre el transporte de mochila y el DE, no son concluyentes. Sin embargo, la noción de que el peso excesivo de las mochilas que llevan niños y adolescentes puede desencadenar un DE, tiene sentido clínico debido, en parte, al aumento de las cargas sobre las estructuras de la columna vertebral [133]. Además, el punto de corte para un “excesivo” peso de la mochila mayor al 10% del peso corporal, está respaldado por la evidencia científica generando consenso entre médicos, investigadores y algunas organizaciones internacionales como la Organización Mundial de la Salud (OMS) [134–136].

Por lo tanto, la evidencia disponible no descarta completamente la posibilidad de que el peso excesivo de las mochilas pueda influir en la presencia, recurrencia o persistencia del DE entre los estudiantes. Sin embargo, cuando se estudie su efecto sería interesante incluir otras variables como el tiempo de transporte o distancia recorrida, la capacidad física del estudiante, entre otras, debido a que es probable que el factor mecánico clave que desencadena el DE cuando se lleva la mochila no sea su peso en sí, sino las circunstancias que lo rodean [133]. Esto explicaría los estudios que muestran que el DL y otras dolencias músculo-esqueléticas son mayores entre los niños que llevan sus mochilas durante períodos más largos, y sugiere que el peso es menos relevante en zonas donde los niños apenas tienen que caminar con sus mochilas que donde tienen que transportarlas durante largos períodos de tiempo [137,138].

1.4. Líneas de acción de la tesis doctoral

Por todo lo mencionado anteriormente, es necesario seguir trabajando para reducir el número de niños y adolescentes con DE, y para poder abordar dicho problema de la forma más adecuada.

De esta forma, los objetivos de la presente tesis se centran en mejorar los conocimientos relacionados con el DE (características del dolor) en escolares y adolescentes de la Región de Murcia (estudio 1), a través de la identificación de los posibles factores asociados a dicho dolor tanto modificables como no modificables (estudio 2). Por otro lado, se centran en la ampliación de los conocimientos relacionados con las pruebas de campos existentes en la actualidad para la valoración de la resistencia de la musculatura del tronco (estudio 3 y 4) y cómo la maduración de los escolares puede influir sobre el desarrollo de la misma (estudio 5). Los principales hallazgos de la presente tesis pretenden ayudar a los profesores de Educación Física y entrenadores, principalmente, además de a todas aquellas personas que trabajen con niños y les preocupe la salud de su espalda, para el desarrollo de adecuadas estrategias de prevención con los escolares.

CAPÍTULO II

OBJETIVOS E HIPÓTESIS

2. Objetivos e hipótesis

2.1. Objetivos generales

Los objetivos generales de esta tesis doctoral se dividieron en dos líneas de investigación. Dentro la primera línea de investigación el objetivo fue profundizar en las características del dolor de espalda en niños y adolescentes, así como los posibles factores asociados al dolor. En la segunda línea de investigación el objetivo fue conocer las principales pruebas de campo que se emplean para la valoración de la resistencia de la musculatura del tronco.

2.2. Objetivos específicos

Los objetivos específicos se han estructurado en función de los cinco estudios de esta tesis doctoral:

Línea de Investigación 1

Estudio 1:

1. Describir la incidencia de niños en edad escolar de la Región de Murcia (España) que sufrieron DE en el año o mes anterior en cuanto a sexo, edad cronológica, etapa de maduración, índice de masa corporal, zona de la espalda afectada, frecuencia, intensidad y severidad del dolor.
2. Determinar los posibles factores asociados con el dolor de espalda.

Estudio 2:

3. Estudiar las características antropométricas y el "Fitness Postural" (disposición sagital integral de la columna vertebral, movilidad de la columna, rango de movimiento de la musculatura flexora y extensora de la cadera y resistencia de los músculos del tronco) en niños y adolescentes con y sin dolor de espalda según el sexo.
4. Determinar si estos factores físicos son posibles desencadenantes de dolor de espalda.

Línea de Investigación 2

Estudio 3:

5. Realizar un meta-análisis de generalización de la fiabilidad para obtener estimaciones de fiabilidad combinadas de las medidas de resistencia de extensión del tronco obtenidas a través de cuatro pruebas de campo.
6. Identificar qué características de los estudios pueden influir en la variabilidad de los coeficientes de fiabilidad.

7. Determinar la práctica de inducción de la fiabilidad en estudios que han utilizado pruebas de resistencia de los músculos extensores del tronco.

Estudio 4:

8. Explorar la fiabilidad inter-sesiones de las medidas obtenidas de 2 pruebas de campo para la valoración de la resistencia de los extensores del tronco (pruebas Biering-Sorensen y Dynamic Extensor Endurance) y 3 pruebas para la valoración de la resistencia de los flexores del tronco (pruebas Ito, Side Bridge y Bench Trunk Curl-Up) en adolescentes de Educación Secundaria Obligatoria.

Estudio 5:

9. Analizar y comparar la influencia de la edad cronológica y la etapa de maduración en la resistencia de los músculos del tronco en chicos y chicas en edad escolar.

2.3. Hipótesis

Las siguientes hipótesis se establecieron en los cinco estudios de esta tesis doctoral:

Línea de Investigación 1

Estudio 1:

1. Presumimos que el dolor de espalda aumentará en frecuencia a medida que avanza la pubertad, así como el peso de los escolares, tanto en niños como en niñas.

Estudio 2:

2. Los resultados del protocolo “Fitness Postural” serán diferentes entre niños y niñas.
3. Los participantes con dolor de espalda presentarán una peor alineación sagital espinal, con un ROM reducido tanto en los flexores como extensores de cadera y una peor resistencia en los músculos del tronco.
4. Los factores físicos asociados al dolor de espalda serán diferentes según el sexo.

Línea de Investigación 2

Estudio 3:

5. La fiabilidad de las medidas obtenidas a través de pruebas de campo para la valoración de la resistencia de los extensores del tronco, va a estar influenciada por las características de los estudios en términos de metodología y características de la muestra.
6. La práctica de inducción de la fiabilidad será común entre los estudios que utilizan pruebas de campo para la valoración de la resistencia de los extensores del tronco.

Estudio 4:

7. A diferencia de las pruebas dinámicas para la valoración de la resistencia de los músculos del tronco, las pruebas isométricas mostrarán una fiabilidad aceptable (para fines clínicos y establecimiento de objetivos) y más sólida que las medidas dinámicas, independientemente del curso de los participantes, debido principalmente a sus sencillos procedimientos de evaluación.

Estudio 5:

8. La resistencia muscular del tronco será similar entre chicos y chicas antes del estirón puberal, y a partir de éste se empezarán a observar diferencias en el rendimiento según el sexo.

CAPÍTULO III

RESULTADOS

3. Resultados

Los resultados que forman parte de la presente tesis doctoral, se presentan a continuación en formato estudios:

Línea de Investigación 1

- ESTUDIO 1: "Prevalence and characteristics of back pain in 9 to 16 years-old children and adolescents from the Region of Murcia (Spain): ISQUIOS Program".
- ESTUDIO 2: "Physical factors associated with back pain in children and adolescents from the Region of Murcia (Spain): ISQUIOS Program".
- ESTUDIO 3: "A Meta-Analysis of the Reliability of Four Field-Based Trunk Extension Endurance Tests".
- ESTUDIO 4: "Reliability of five trunk flexion and extension endurance field-based tests in high school-aged adolescents: ISQUIOS Program".
- ESTUDIO 5: "Effect of maturation on trunk muscle endurance in 8 to 16-year-old children and adolescents from the Region of Murcia (Spain): ISQUIOS Program".

Además, de estos estudios, el proceso de la tesis doctoral ha dado como resultado otras publicaciones relacionadas directamente con la presente tesis entre las que destacan:

- Santonja-Medina F, Collazo-Diéguez M, Martínez-Romero MT, Rodríguez-Ferrán O, Aparicio-Sarmiento A, Cejudo A, Andújar P, Sainz de Baranda P. Classification System of the Sagittal Integral Morphotype in Children from the ISQUIOS Programme (Spain). Int J Environ Res Public Health. 2020 Apr 4;17(7):2467. doi: 10.3390/ijerph17072467. PMID: 32260344; PMCID: PMC7177434.
- Sainz de Baranda P, Andújar P, Collazo-Diéguez M, Pastor A, Santonja-Renedo F, Martínez-Romero MT, Aparicio-Sarmiento A, Cejudo A, Rodríguez-Ferrán O, Santonja-Medina F. Sagittal standing spinal alignment and back pain in 8 to 12-year-old children from the Region of Murcia, Spain: The ISQUIOS Program. J Back Musculoskelet Rehabil. 2020 Aug 27. doi: 10.3233/BMR-191727. Epub ahead of print. PMID: 32924979.
- Sainz de Baranda, P., Cejudo, A., Martínez-Romero, M. T., Aparicio-Sarmiento, A., Rodríguez-Ferrán, O., Collazo-Diéguez, M., Hurtado-Avilés, J., Andújar, P., & Santonja-Medina, F. (2020). Sitting Posture, Sagittal Spinal Curvatures and Back Pain in 8 to 12-Year-Old Children from the Region of Murcia (Spain): ISQUIOS Programme. International journal of environmental research and public health, 17(7), 2578. https://doi.org/10.3390/ijerph17072578

- Aparicio-Sarmiento, A.; Rodríguez-Ferrán, O.; Martínez-Romero, M.T.; Cejudo, A.; Santonja, F.; Sainz de Baranda, P. Back Pain and Knowledge of Back Care Related to Physical Activity in 12 to 17 Year Old Adolescents from the Region of Murcia (Spain): ISQUIOS Programme. Sustainability 2019, 11, 5249.
- Sainz de Baranda P, Cejudo A, Moreno-Alcaraz VJ, Martinez-Romero MT, Aparicio-Sarmiento A, Santonja-Medina F. Sagittal spinal morphotype assessment in 8 to 15 years old Inline Hockey players. PeerJ. 2020 Jan 2;8:e8229. doi: 10.7717/peerj.8229. PMID: 31915572; PMCID: PMC6942677.
- Ginés-Díaz A, Martínez-Romero MT, Cejudo A, Aparicio-Sarmiento A, Sainz de Baranda P. Sagittal Spinal Morphotype Assessment in Dressage and Show Jumping Riders. J Sport Rehabil. 2019 Oct 18:1-8. doi: 10.1123/jsr.2018-0247. Epub ahead of print. PMID: 31034307.
- Hernández-García R, Gil-López MI, Martínez-Pozo D, Martínez-Romero MT, Aparicio-Sarmiento A, Cejudo A, Baranda PS, Bishop C. Validity and Reliability of the New Basic Functional Assessment Protocol (BFA). Int J Environ Res Public Health. 2020 Jul 5;17(13):4845. doi: 10.3390/ijerph17134845. PMID: 32635678; PMCID: PMC7369969.
- Aparicio-Sarmiento A, Martínez-Romero MT, Rivas, A., Rodríguez-Ferrán, O. (2020). Características de las intervenciones y estudios sobre el cuidado de la espalda en la etapa escolar. JUMP, 2, 37-47. https://doi.org/10.17561/jump.n2.4
- Aparicio-Sarmiento, A.; Rivas, A.; Martínez-Romero, M.T.; Rodríguez-Ferrán, O.; Sainz de Baranda, P. (2018). La valoración del conocimiento sobre educación postural en Educación Primaria a través de la Batería ISQUIOS: La opinión del docente. Revista Pedagógica ADAL, 21(36), 13-20.

Además, parte del trabajo del proceso desarrollado a lo largo de estos años ha quedado plasmado en algunos capítulos del Libro: Programa Educativo ISQUIOS. Investigación y aplicaciones prácticas en el contexto escolar (ISBN: 978-84-09-17399-0).

ESTUDIO 1: Prevalence and characteristics of back pain in 9 to 16 years-old children and adolescents from the Region of Murcia (Spain): ISQUIOS Program

3.1. Introduction

Today, back pain (BP) during childhood and adolescence is no longer considered uncommon or rare [1,2]. Most of the studies focus on investigating the prevalence of low back pain (LBP) because it is the most common among children and adolescents [3]. However, the pain in the neck and thoracic area (mid-back) should not be underestimated because the prevalence is becoming more frequent at this growth stage [4,5].

Aartun et al., (2014) found that neck pain (NP), mid-back pain (MBP) and LBP were common at the age of 11-13 years, and two years later this prevalence increased for each area. Regarding the course of pain, they concluded that for the majority of the participants, the pain seems to be mild in nature, relatively infrequent and of low intensity, and a small group (14-20%) was more severely affected with frequent pain which was also more intense [6]. In a Spanish sample, BP prevalence within the last week for school-children (8-12 years) was reported to be 10.6% (1.7% NP, 7.7 MBP and 2.9% LBP) and for high-school students (12-17 years) the BP prevalence in the last week was 25.7% [7,8]. The presence of BP during childhood and adolescence might generate school absenteeism, the need for medical attention and medication [2,9], along with activity of daily living (ADL) limitations (i.e. standing in a queue, carrying a backpack or doing physical activities) [1,2].

Given that the prevalence of BP increases from childhood to adolescence, more specifically at around 11-12 years or even more after the onset of puberty [10–12], it appears that puberty (transition period characterized by rapid physiological and anatomical changes) is the time of a rapid increase in the prevalence values of BP, being this prevalence higher among girls due to their early onset of pubertal development (an average of 2 years earlier than boys) [3,11,13]. A recent prospective study of 3 years about spinal pain in children (9 years) concluded that both pubertal development and linear growth were associated with spinal pain. It was observed that boys and girls with more advanced pubertal development and those undergoing greater growth experienced increased spinal pain frequency and duration [14].

It is difficult to understand that even knowing the influence of puberty on the appearance of BP, research is still being carried out on the BP prevalence based on the school year or the chronological age (CA) of children and adolescents. There is a large individual variability in the timing and tempo of the maturational process, resulting in differences in size, shape and body composition in children and adolescents of similar age [15–17]. At a particular age, adolescents of the same sex can be in very different stages of maturation, some who have not started pubertal development and others who have fully completed it. Therefore, CA is not an appropriate reference when working with young people [17].

It is well known that, during the peak height velocity (PHV), musculoskeletal structures of children are not mature enough to support the sudden mechanical loading changes on the spine generated by the differential growth rates that exist between the legs and the trunk, the long bones of the legs experience a peak growth in front of the shorter bones of the trunk [18,19]. Furthermore, muscle development and the adaptation of ligaments and tendons to these new loads do not follow a similar onset and rate either. Therefore, the phase of growth spurt could be considered a critical period compared with the episodes before or after that stage [11,14,20,21].

These structural changes, along with physical (increased height and change in body composition) and hormonal changes (females secreting more estrogen and males secreting more testosterone) [13,18,22] lead to a greater height and muscle mass in boys and increased body fat in girls during the onset of puberty [23–25]. Weight gain during maturation in girls has been associated with increased accumulation of fat mass, which can be related to low competence in motor skills and low participation in physical activities [15,16]. It has been proposed that overweight and obese may influence the appearance of BP [2,26,27]. Furthermore, the mechanical load on the spine can increase in people with obesity due to the increased compression force on the structures of the lumbar spine [26]. In this context, childhood obesity may represent a modifiable risk factor for back pain, and weight optimisation confers benefits beyond musculoskeletal health [2,27].

The puberty is a vulnerable period for musculoskeletal disorders due to the existence of a wide inter-individual variation in growth and development. Therefore, the aims of this study were a) to describe the prevalence of school-aged children from the Region of Murcia (Spain) who suffered from BP within the previous year or month in terms of sex, CA, maturational stage, BMI, affected region of the back, frequency, intensity, and severity of pain, and b) to determine the possible factors associated with this disorder. We hypothesize that BP will increase in prevalence as puberty progresses, as well as the weight of schoolchildren, both in boys and girls.

3.2. Method

3.2.1. Design

A cross-sectional study design was used to identify BP prevalence in children and adolescents of different sex, CA and maturity status. All measurements were collected before participate in a postural and physical fitness program called “ISQUIOS Program”. The study was conducted during the first term of the 2017–2018 and 2018-2019 school years.

3.2.2. Participants

A total of 823 students (primary education from 8 to 12 years old, n=541; secondary education from 12 to 17 years old, n=282) were initially invited from 10 elementary schools and 2 Secondary schools of the Region of Murcia (Spain) to participate in this study (convenience sample). The exclusion criteria were: a) to have a diagnosed spinal pathology or important physical injuries, b) do not have returned the signed informed consent (both from parents/guardians and students) before the start of the study, and c) do not have completed the BP questionnaire.

A comprehensive verbal description of the nature and purpose of the study was given to the students and their parents/guardians and PE teachers. The protocol was fully approved by the Review Committee for Research Involving Human Subjects at the University of Murcia (Spain) (ID: 1920/2018) and according to the Declaration of Helsinki.

Finally, a sample of 513 school and high school students from 9 to 16 years old (49.9% females; mean age ±SD = 12.6±1.9 years) completed this study. One hundred and fifteen students were removed based on the exclusion criteria (Figure 2).

3.2.3. Procedure

Participants were tested during two different sessions within PE classes. In the first session, the anthropometric measures were recorded to calculate the maturation stage of the students. In the second session, an experienced researcher presented the questionnaire to the older students (secondary school) and the parents (in case of primary school students), explained the procedure to complete the survey and personally solved all participant questions. Secondary school students filled out the questionnaire during PE class, whereas primary school children filled out the BP questionnaire with their parents at home. The use of parental reports is very important with early school-age children (under 11 years old) because it helps to improve the quality of the information gathered and has been recommended in prior studies [28].

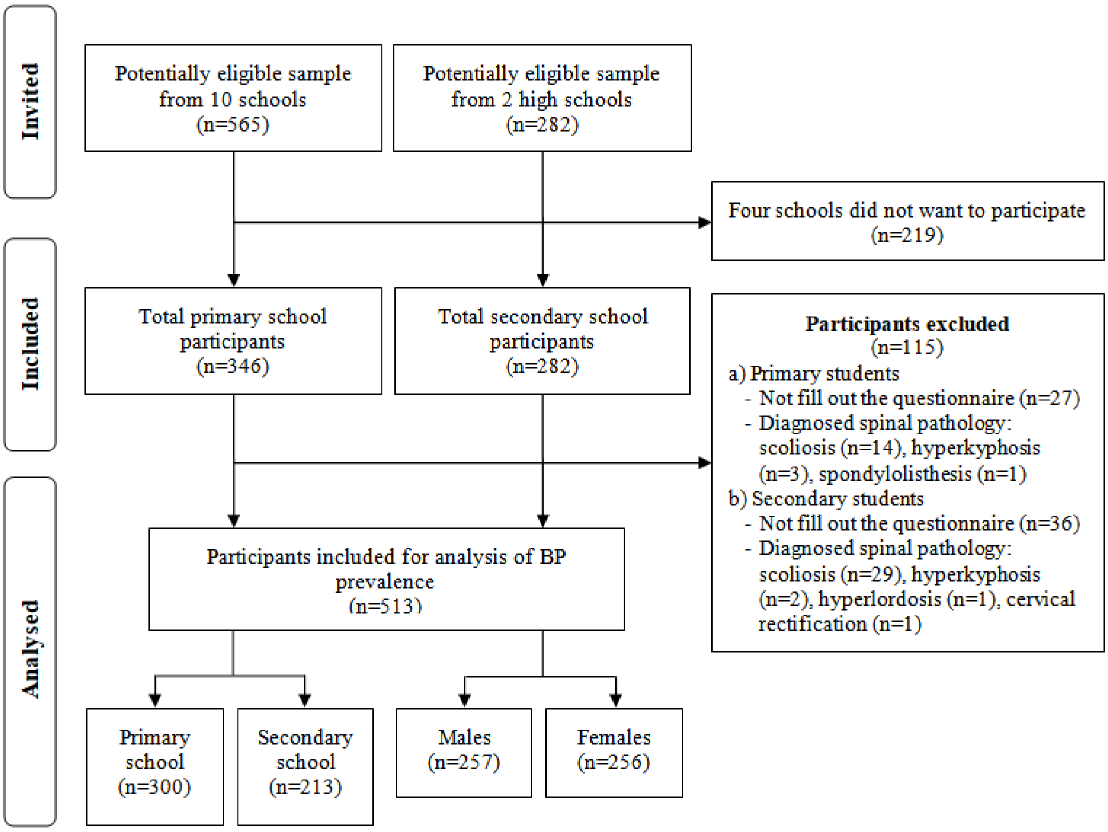


Figure 2. Flow diagram for the sample selection

Anthropometry and stage of maturation

Body mass in kilograms was measured on a calibrated physician scale (SECA 799, Hamburg, Germany). Body height was recorded in centimetres on a measurement platform (SECA 799). Sitting height was measured in centimetres. Leg length was calculated as the difference between body height and sitting height. Body mass index (BMI) was obtained by dividing body mass by height (in meters) squared and was classified according to the US Centres for Disease Control and Prevention's age- and sex-matched percentile grading: underweight (equal to or below 5%), normal weight (from 5.01% to 84.99%), overweight (from 85% to 94.99%), and obese (equal to or above 95%).

Stage of maturation was calculated in a non-invasive manner using a regression equation comprising measures of age, body mass, body height, sitting height and leg length taken during the first part of the testing sessions [29]. Using this method, maturity offset (calculation of years from PHV) was completed (equation 1 for boys and equation 2 for girls). Due to the error in the prediction equation of approximately 6 months in the paediatric population [29], participants with a maturity offset of -0.99 to -0.51 years and +0.51 to +0.99 years were removed from the analysis by maturational stage [30,31]. Also, participants whose maturity offset were outside −3 or +3 years, were removed from the analysis by maturational stage to maximize accuracy [30]. This approach enabled the identification of 3 different maturity groups: pre-PHV (maturity offset of less than −1), circa-PHV (maturity offset between −0.5 and +0.5), and post-PHV (maturity offset of greater than +1). Therefore, the following equations to calculate maturity offset were used:

Equation 1 (boys) = $-9.236 + 0.0002708 \times (Leg\ Length \times Sitting\ Height) - 0.001663 \times (Age\ \times Leg\ Length) + 0.007216 \times (Age \times Sitting\ Height) + 0.02292 \times (Weight/Height \times 100)$

Equation 2 (girls) = $-9.376 + 0.0001882 \times (Leg\ Length \times Sitting\ Height) + 0.0022 \times (Age\ \times Leg\ Length) + 0.005841 \times (Age \times Sitting\ Height) - 0.002658 \times (Age \times Weight) + 0.07693 \times (Weight/Height \times 100)$

BP Assessment

An ad-hoc questionnaire composed of 8 items was used to describe BP prevalence in school-age males and females [3,7,8,32,33]. The reliability and validity of the questionnaire have been evaluated in previous studies with Spanish children and adolescents [34,35].

The questionnaire starts with sociodemographic issues such as sex, age, school, grade level, and pathologies diagnosed. After that, there are questions about BP prevalence within the last 12 months and last 1-month ("During the last year (or last month), have you had back pain?"). Those who experienced BP within the last year completed the items based on the spinal pain areas (neck, mid-back and/or low back) and the characteristics of pain within the last year (frequency and intensity of pain, irradiation to lower limb (sciatica) and limitations of daily life activities). BP was defined as aching, pain, or discomfort in some parts of the back that were not related to trauma or menstrual pain. The questionnaire included a drawing of the back to mark the most affected area (Figure 3). The frequency of spinal pain was classified as "at least once", "some days", "often (several days)" or "daily". The intensity of spinal pain was quantified using a visual analogue scale (VAS) ranging from 0 (no pain) to 10 (intolerable pain).

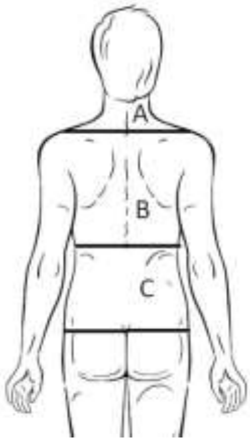


Figure 3. Drawing of the back to mark the area of back pain.
A: neck pain, B: mid-back pain, C: low back pain

3.2.4. Statistical Analyses

Descriptive statistics were performed for the total sample, as well as by sex, CA and maturational stage including mean values and standard deviations (SD) for the quantitative variables, and counts and percentages for qualitative variables.

Chi-square tests (bivariate analysis) were applied to explore the associations between the presence of BP (dependent variable) with sex, CA groups, maturational groups and BMI classifications (or Fisher's exact test). Due to the small numbers and to satisfy the requirements of applicability of the chi-square independence test, the variable BMI was divided into underweight/normal weight and overweight/obesity. Any significant associations by the bivariate analysis were furthered analysed by binary logistic regression to identify the relationship of the dependent variable with independent variables and to confirm the results from the analyses on the chi-square test. The resulting crude odds ratio (OR), only one explanatory variable included in the regression, and associated 95% confidence intervals (95%CI) were reported. Effect sizes for the OR were defined as follows: small effect OR = 1 to 1.25, medium effect OR = 1.25 to 2 and large effect OR ≥ 2. The independent t-test was used to explore the differences between students with or without BP (last year and last month) and anthropometric characteristics. It was also used to analyse the differences between those who reported ADL limitation and pain intensity. For exploring the differences between pain intensity and pain localization or frequency, one-way ANOVA was conducted with post hoc Scheffe test.

The statistical analysis was conducted using the statistical package SPSS v. 21.0 for Windows (IBM Corp., Armonk, NY, USA).

3.3. Results

This study was based on a sample of 513 students, aged between 9 and 16 years (12.6±1.9 years); 257 (50.1%) were males and 256 (49.9%) females. Descriptive statistics for each CA groups and maturation stage are displayed in Table 5.

Table 5. Descriptive anthropometric values (mean ± SD) for participants per chronological age groups and per maturational status adjusted by sex.

CA group	Sex	N	Age (y)	Body mass (kg)	Body height (cm)	BMI	BMI percentile	Maturity offset
9	M	38	9.6±0.2	40.6±11.1	140.9±6.1*	20.1±4.3	72.9±30.9	-2.7±0.4*
	F	38	9.5±0.2	36.3±8.4	136.4±7.4	19.3±3.1	72.1±27.9	-2.1±0.4
10	M	47	10.5±0.2	40.1±8.5	143.7±6.6	19.4±3.4	66.5±30.2	-2.6±0.3*
	F	59	10.4±0.2	42.4±12.1	142.8±7.3	20.5±4.5	68.8±29.9	-1.3±0.6
11	M	47	11.5±0.3	45.2±11.2	147.1±6.9*	20.4±3.9	70.1±28.2	-2±0.4*
	F	54	11.6±0.3	47.4±10.3	150.9±7.1	20.9±4.1	67.5±27.3	-0.2±0.5
12	M	39	12.4±0.3	52.8±13.1	155.5±9.3	21.4±4.1	71.3±30.2	-1.1±0.6*
	F	33	12.4±0.3	49.8±12	153.5±5.3	21.1±4.3	65.6±29.3	0.4±0.5
13	M	30	13.4±0.2	57.7±13.1	162.6±9.4*	22±4	70.5±26.6	-0.3±0.8*
	F	19	13.4±0.3	59.3±13.5	156.2±4.8	24±5	77.6±25.3	1.2±0.5
14	M	37	14.6±0.2	63.3±11.3	167.6±6.5*	22.6±3.9	68.4±25.2	0.6±0.6*
	F	30	14.5±0.2	58±12.1	160.6±4.6	22.8±4.6	64.7±30.7	2±0.3
>15	M	19	15.9±0.7	60.6±13.1	169.9±7.4*	20.8±3.9	46.5±34.4*	1.4±0.6*
	F	23	15.7±0.6	59.3±9	160.1±6.3	22.8±3.4	70.9±20.5	2.5±0.4
Maturity status								
Pre-PHV	M	142	11.2±1*	44.3±10.1*	146.5±6.7*	20.4±3.8*	71±29.8	-2.2±0.6
	F	90	10.1±5	37.6±9.1	138.6±6.4	19.4±3.7	66.7±29.1	-1.8±0.5
Circa-PHV	M	57	13.8±0.9*	60.3±11.2*	164.8±6.2*	22.3±3.9	69.3±26.7	0.04±0.4*
	F	81	11.7±0.6	49.3±9.5	152.4±5.1	21.2±4	69.9±28.3	-0.02±0.3
Post-PHV	M	27	15.3±0.8*	66±13.5*	172.6±5.2*	22.3±4.6	59.2±33.9	1.4±0.4
	F	69	14.5±1	59.9±11.5	159.7±4.4	23.4±4.4	71.8±26.8	1.9±0.5

Note: M: males, F: females, y: years, kg: kilograms, cm: centimetres, BMI: body mass index, CA: chronological age, PHV: peak height velocity. *Significant differences between boys and girls of the same age or in the same maturational stage

Among the main characteristics, it should be noted that 232 students (49.8%) were in a state before the onset of puberty, 138 participants (39.6%) were in the pubertal stage, and 96 (20.6%) had already passed the puberty. On the other hand, concerning the BMI classification, 219 scholars (42.7%) are overweight or obese and 294 students (57.3%) are underweight or normal weight. According to CA, the main differences are found in the height of the participants, the height being higher in boys at all ages except 11 years, and in the maturity offset, being always higher among girls. Regarding the maturational state, differences are found mainly between boys and girls in age, height and weight.

3.3.1. Back pain prevalence in the last year and last month

Table 6 shows the counts and frequencies of BP in the last year and last month for the total sample and the variables: sex, CA group, maturity status and BMI classification. The results of the associations between the presence of BP and all the variables mentioned were also presented. The presence of BP in the last year was observed in 180 (35.1%) students, 81 were boys (45%) and 99 were girls (55%), no association was found between BP and sex. For the 1-month prevalence, 89 (17.3%) students presented BP, of which 36 were boys (40.5%) and 53 were girls (59.5%), finding an association between having BP the previous month and being a girl. The crude OR from the binary logistic regression confirmed the results from the chi-square for BP in the last month and female sex (OR: 1.60, 95% CI: 1.01-2.55).

According to the CA of the participants, the prevalence of BP in the last year and last month grew as the students' age increased (last year: from 27.6% among participants aged 9 to 57.1% among students over 15 years, last month: from 10.5% among participants aged 9 to 30.9% among students over 15 years). It was found that there were associations between BP prevalence in the last year and CA groups and between BP in the last month and CA groups. The crude OR from the binary logistic regression confirmed the results from the chi-square for BP in the last year and 12- (OR: 2.34, 95% CI: 1.18-4.64), and 15-year-old (OR: 3.49, 95% CI: 1.59-7.7). As for BP in the last month, it was associated with 12- (OR: 3.05, 95% CI: 1.24-7.50), 14- (OR: 2.89, 95% CI: 1.16-7.22) and 15-year-old (OR: 3.81, 95% CI: 1.43-10.17).

The same trend appeared when observing the prevalence of BP in the last year and las month according to the maturational stage. Among the participants in the pre-PHV state was found a 1-year prevalence of 28.4% (last month: 12.1), within the circa-PHV state it increased to 34.1% (last month: 14.5) and in the post-PHV state, the prevalence was 52.1% (last month: 34.4%). Associations were found between post-PHV state and having BP in the last year and last month. The crude OR from the binary logistic regression confirmed the results from the chi-square for BP in the last year and post-PHV (OR: 2.73, 95% CI: 1.67-4.47) and for BP in the last month and post-PHV (OR: 3.81, 95% CI: 2.14-6.80).

Table 6. Prevalence of back pain during the last year and last month for the total sample and its association with sex, chronological age group, maturity status and body mass index classification.

Variables	Back pain last year No (333)	Back pain last year Yes (180)	Back pain last year Chi-square	Back pain last month No (424)	Back pain last month Yes (89)	Back pain last month Chi-square
Sex						
M	52.9% (176)	45% (81)	$\chi^2_{(1)}$=2.882 p=0.09	52.1% (221)	40.4% (36)	$\chi^2_{(1)}$=4.009 p=0.045
F	47.1% (157)	55% (99)		47.9% (203)	59.6% (53)	
CA group						
9	16.5% (55)	11.7% (21)		16% (68)	9% (8)	
10	23.7% (79)	15% (27)		22.6% (96)	11.2% (10)	
11	21.3% (71)	16.7% (30)		21% (89)	13.5% (12)	
12	11.4% (38)	18.9% (34)	$\chi^2_{(6)}$=21.791 p=0.001	12.5% (53)	21.3% (19)	$\chi^2_{(6)}$=22.056 p=0.001
13	8.7% (29)	11.1% (20)		9.2% (39)	11.2% (10)	
14	12.9% (43)	13.3% (24)		11.8% (50)	19.1% (17)	
>15	5.4% (18)	13.3% (24)		6.8% (29)	14.6% (13)	
Maturity status						
Pre-PHV	54.8% (166)	40.5% (66)		53% (204)	34.6% (28)	
Circa-PHV	30% (91)	28.8% (47)	$\chi^2_{(2)}$=16.751 p<0.001	30.6% (118)	24.7% (20)	$\chi^2_{(2)}$=24.666 p<0.001
Post-PHV	15.2% (46)	30.7% (50)		16.4% (63)	40.7% (33)	
BMI classification						
Underweight/ Normal	61.3% (204)	50% (90)	$\chi^2_{(1)}$=6.05 p=0.01	59.2% (251)	48.3% (43)	$\chi^2_{(1)}$=3.561 p=0.059
Overweight/ Obese	38.7% (129)	50% (90)		40.8% (173)	51.7% (46)	

Note: M: males, F: female, CA: chronological age, PHV: peak height velocity, BMI: body mass index

According to the BMI classification, the prevalence of BP also increased from those with low or normal weight (last year: 30.6%, last month: 14.6%) to those with overweight or obesity (last year: 41.1%, last month: 21%), finding an association between being overweight or obese and presenting BP in the last year. The crude OR from the binary logistic regression confirmed the results from the chi-square for BP in the last year and overweight/obese status (OR: 1.58, 95% CI: 1.09-2.28).

Comparing the anthropometric characteristics of the participants between those who had BP during the previous year or last month and those who did not, students with BP were older, taller, and had a greater trunk length, leg length, body mass and BMI ($p<0.001$).

Table 7 shows the counts and frequencies of BP in the last year and last month adjusted by sex. The results of the associations between the presence of BP and all the variables were also presented.

In boys, there were associations between the CA group and BP in the last year and month. An association was also found between having BP during the past year and BMI. The crude OR values from the binary logistic regression confirmed the association between BP and 12-year-old (OR: 2.58, 95% CI: 1.01-6.62) and between BP and overweight/obese status (OR: 1.14, 95% CI: 1.14-3.30). The association between BP in the last month and CA was confirmed through regression for the 12-year-old group (OR: 4.58, 95% CI: 1.16-18.03).

On the other hand, in girls, associations were found between BP in the last year and the last month and the CA groups, as well as with the maturation stage. The binary logistic regression confirmed the association between having BP during the past year and being 13- (OR: 6.07, 95% CI: 1.81-20.28), 14- (OR: 2.80, 95% CI: 1.81-7.34) or 15-years-old (OR: 5.25, 95% CI: 1.71-16.11) and being in a post-PHV state (OR: 3.79, 95% CI: 1.93-7.41). For BP during the previous month, the associations were with 13- (OR: 4.80, 95% CI: 1.29-17.77), 14-year-old (OR: 4.40, 95% CI: 1.33-14.47) and post-PHV (OR: 5.44, 95% CI: 2.34-12.68).

Table 7. Prevalence of back pain during the last year and last month adjusted by sex and its association with chronological age group, maturity status and body mass index classification.

Variables	Back pain last year						Back pain last month					
	Male			Female			Male			Female		
	No (176)	Yes (81)	Chi-square	No (157)	Yes (99)	Chi-square	No (221)	Yes (36)	Chi-square	No (203)	Yes (53)	Chi-square
CA group												
9	15.3% (27)	13.6% (11)		17.8% (28)	10.1% (10)		15.8% (35)	8.3% (3)		16.3% (33)	9.4% (5)	
10	19.3% (34)	16% (13)		28.7% (45)	14.1% (14)		19% (42)	13.9% (5)		26.6% (54)	9.4% (5)	
11	19.9% (35)	14.8% (12)		22.9% (36)	18.2% (18)		19% (42)	13.9% (5)		23.2% (47)	13.2% (7)	
12	10.8% (19)	24.7% (20)	$\chi^2_{(6)}$=12.309 *p*=0.05	12.1% (19)	14.1% (14)	$\chi^2_{(6)}$=24.39 *p*<0.001	12.7% (28)	30.6% (11)	$\chi^2_{(6)}$=12.330 *p*=0.05	12.3% (25)	15.1% (8)	$\chi^2_{(6)}$=23.797 *p*=0.001
13	13.1% (23)	8.6% (7)		3.8% (6)	13.1% (13)		12.7% (28)	5.6% (2)		5.4% (11)	15.1% (8)	
14	15.9% (28)	11.1% (9)		9.6% (15)	15.2% (15)		14.5% (32)	13.9% (5)		8.9% (18)	22.6% (12)	
>15	5.7% (10)	11.1% (9)		5.1% (8)	15.2% (15)		6.3% (14)	13.9% (5)		7.4% (15)	15.1% (8)	
Maturity status												
Pre-PHV	64.3% (99)	59.7% (43)		45% (67)	25.3% (23)		63.7% (123)	57.6% (19)		42.2% (81)	18.8% (9)	
Circa-PHV	25.3% (39)	25% (18)	$\chi^2_{(2)}$=1.146 *p* =0.56	34.9% (52)	31.9% (29)	$\chi^2_{(2)}$=16.142 *p*<0.001	25.9% (50)	21.2% (7)	$\chi^2_{(2)}$=3.193 *p*=0.20	35.4% (68)	27.1% (13)	$\chi^2_{(2)}$=19.897 *p*<0.001
Post-PHV	10.4% (16)	15.3% (11)		20.1% (30)	42.9% (39)		10.4% (20)	21.2% (7)		22.4% (43)	54.2% (26)	
BMI classification												
Underweight/ Normal	61.9% (109)	45.7% (37)	$\chi^2_{(1)}$=5.517	60.5% (95)	53.5% (53)	$\chi^2_{(1)}$=6.927	58.4% (129)	47.2% (17)	$\chi^2_{(1)}$=4.233	60.1% (122)	49.1% (26)	$\chi^2_{(1)}$=3.153
Overweight/ Obese	38.1% (67)	54.3% (44)	*p*=0.02	39.5% (59)	46.5% (44)	*p*=0.07	41.6% (92)	52.8% (19)	*p*=0.24	39.9% (81)	50.9% (27)	*p*=0.37

Note: CA: chronological age, PHV: peak height velocity, BMI: body mass index

3.3.2. Prevalence of back pain by spinal area in the last year

The spinal area where the pain was most often reported was the lower back (59.4%), followed by the mid-back (37.2%) and neck area (21.1%). During the last year, 13.9% of students reported pain in more than one spinal area. Table 8 shows the counts and frequencies of BP in the last year according to reported spinal pain area for the total sample and the variables: sex, CA group, maturity status and BMI classification. The results of the associations between the presence of BP in any spinal area and all the variables mentioned were also presented.

According to the sex of the participants, girls showed a higher prevalence of BP (NP: 8.2%, MBP: 16.4% and LBP: 21.5%) for all spinal areas than boys (NP: 6.6%, MBP: 9.7% and LBP: 20.23%), although a significant association was only found between sex and MBP. The crude OR from the binary logistic regression confirmed the results from the chi-square for MBP and female sex (OR: 1.82, 95% CI: 1.07-3.09).

For CA groups, the prevalence in the different spine areas increased with age (NP: from 3.9% to 11.9%, MBP: from 11.8% to 19.05%, LBP: from 18.4% to 40.5%), finding a significant association between CA and NP and LBP. The crude OR from the binary logistic regression confirmed the results from the chi-square for NP and 13-year-old (OR: 5.47, 95% CI: 1.40-21.38) and for LBP and 15-year-old (OR: 3.01, 95% CI: 1.29-7.02).

Regarding the maturity status, there was also an increase in the prevalence of pain according to the pubertal development of the participants, with the highest increase in the post-PHV state (NP: from 6.1% to 12.5%, MBP: from 10.8 to 21.9, LBP: from 18.5 to 28.1). A significant association was found between being in a post-PHV state and having MBP. The crude OR from the binary logistic regression confirmed the results from the chi-square for MBP and post-PHV (OR: 2.32, 95% CI: 1.23-4.38).

Table 8. Prevalence of back pain during the last year according to reported spinal pain area for the total sample and its association with sex, chronological age group, maturity status and body mass index classification.

Variables	Neck pain No (475)	Neck pain Yes (38)	Neck pain Chi-square	Mid-back pain No (446)	Mid-back pain Yes (67)	Mid-back pain Chi-square	Low back pain No (424)	Low back pain Yes (89)	Low back pain Chi-square
Sex									
M	50.5% (240)	44.7% (17)	$\chi^2_{(1)}$=0.472 *p*=0.492	52% (232)	37.3% (25)	$\chi^2_{(1)}$=5.038 *p*=0.025	50.5% (205)	48.6% (52)	$\chi^2_{(1)}$=0.1223 *p*=0.727
F	49.5% (235)	55.3% (21)		48% (214)	62.7% (42)		49.5% (201)	51.4% (55)	
CA group									
9	15.4% (73)	7.9% (3)		15% (67)	13.4% (9)		15.3% (62)	13.1% (14)	
10	21.1% (100)	15.8% (6)		21.3% (95)	16.4% (11)		22.4% (91)	14% (15)	
11	20.2% (96)	13.2% (5)		20.4% (91)	14.9% (10)		19.5% (79)	20.6% (22)	
12	13.5% (64)	21.1% (8)	$\chi^2_{(6)}$=13.867 *p*=0.02	14.1% (63)	13.4% (9)	$\chi^2_{(6)}$=5.444 *p*=0.488	12.6% (51)	19.6% (21)	$\chi^2_{(6)}$=20.802 *p*=0.002
13	8.4% (40)	23.7% (9)		9.4% (42)	10.4% (7)		11.1% (45)	3.7% (4)	
14	13.7% (65)	5.3% (2)		12.1% (54)	19.4% (13)		13.1% (53)	13.1% (14)	
>15	7.8% (37)	13.2% (5)		7.6% (34)	11.9% (8)		6.2% (25)	15.9% (17)	
Maturity status									
Pre-PHV	50.6% (218)	40% (14)		51.2% (207)	40.3% (25)		51.1% (112)	44.8% (43)	
Circa-PHV	29.9% (129)	25.7% (9)	$\chi^2_{(2)}$=4.362 *p*=0.113	30.2% (122)	25.8% (16)	$\chi^2_{(2)}$=7.749 *p*=0.02	30.3% (112)	27.1% (26)	$\chi^2_{(2)}$=4.190 *p*=0.123
Post-PHV	19.5% (84)	34.3% (12)		18.6% (75)	33.9% (21)		18.6% (69)	28.1% (27)	
BMI classification									
Underweight/Normal	58.3% (277)	44.7% (17)	$\chi^2_{(1)}$=2.652 *p*=0.103	57.2% (255)	58.2% (39)	$\chi^2_{(1)}$=0.025 *p*=0.873	60.3% (245)	45.8% (49)	$\chi^2_{(1)}$=7.328 *p*=0.007
Overweight/Obese	41.7% (198)	55.3% (21)		42.8% (191)	41.8% (28)		39.7% (161)	54.2% (58)	

Note: M: male, F: female, CA: chronological age, PHV: peak height velocity, BMI: body mass index

According to BMI classification, the prevalence of LBP in overweight or obese participants (26.5%) was the only one that increased considerably compared to their normal-weight peers (16.7%), finding an association between being overweight or obese and presenting LBP. The crude OR from the binary logistic regression confirmed the results from the chi-square for LBP and overweight/obese status (OR: 2.87, 95% CI: 1.49-5.55)

No differences or associations were found between those who had pain in more than one spinal area and the anthropometric or sociodemographic characteristics of the students.

When analysing the possible associations between having pain in one area or another of the spine, an association was found between those with NP and those with MBP ($\chi^2_{(1)}$=12.395, p<0.001, OR: 3.52, 95% CI: 1.68-7.83).

Adjusted by sex, the prevalence of BP in the last year in girls according to the area of spinal pain showed an association between NP and CA groups ($\chi^2_{(6)}$=15.608, p=0.006) and maturational stage ($\chi^2_{(2)}$= 7.570, p=0.01). The association between NP and CA was confirmed through regression for 13- (OR: 13.21, 95% CI: 1.42-123.31) and 15-year-old (OR: 10.28, 95% CI: 1.17-94.61), and between NP and maturational status for post-PHV (OR: 4.08, 95% CI: 1.23-13.43). For LBP in girls, an association was also found with CA ($\chi^2_{(6)}$=16.843, p=0.01) and maturational status ($\chi^2_{(2)}$=6.141, p=0.04), in this case, the regression showed an association with 15 years (OR: 5.82, 95% CI: 1.76-19.23) and post-PHV (OR: 2.59, 95% CI: 1.19-5.65).

In boys, the only association was found between LBP and CA groups ($\chi^2_{(6)}$=15.021, p=0.02) and BMI classification ($\chi^2_{(1)}$=4.204, p=0.04). The crude OR from the binary logistic regression confirmed the results from the chi-square for LBP and post-PHV (OR: 1.89, 95% CI: 1.02-3.49).

3.3.3. Prevalence of back pain by spinal area in the last month

Concerning the areas of spinal pain during the previous month, practically the same trend appeared as during the past year. The spinal pain area most often reported during the last month was the lower back (58.4%), followed by the mid-back (34.8%) and neck area (23.6%). During the last month, 15.7% reported pain in more than one spinal area. Table 9 shows the counts and frequencies of BP in the last month according to reported spinal pain area for the total sample and the variables: sex, CA group, maturity status and BMI classification. The results of the associations between the presence of BP in any spinal area and all the variables mentioned were also presented.

Table 9. Prevalence of back pain during the last month according to reported spinal pain area for the total sample and its association with sex, chronological age group, maturity status and body mass index classification.

Variables	Neck pain			Mid-back pain			Low back pain		
	No (492)	Yes (21)	Chi-square	No (482)	Yes (31)	Chi-square	No (461)	Yes (52)	Chi-square
Sex									
M	50.4% (248)	42.9% (9)	$\chi^2_{(1)}=0.459$ $p=0.498$	50.4% (243)	45.2% (14)	$\chi^2_{(1)}=0.322$ $p=0.571$	50.5% (238)	48.6% (19)	$\chi^2_{(1)}=4.255$ $p=0.04$
F	49.6% (244)	57.1% (12)		49.6% (239)	54.8% (17)		49.5% (223)	51.4% (33)	
CA group									
9	15% (74)	9.5% (2)	$\chi^2_{(6)}=6.581$ $p=0.318$	14.9% (72)	12.9% (4)	$\chi^2_{(6)}=4.100$ $p=0.661$	15.6% (72)	7.7% (4)	$\chi^2_{(6)}=21.799$ $p=0.001$
10	20.9% (103)	14.3% (3)		21% (101)	16.1% (5)		22.1% (102)	7.7% (4)	
11	20.9% (98)	14.3% (3)		20.1% (97)	12.9% (4)		20.2% (93)	15.4% (8)	
12	13.6% (67)	23.8% (5)		13.9% (67)	16.1% (5)		13% (60)	23.1% (12)	
13	9.1% (45)	19% (4)		9.5% (46)	9.7% (3)		10% (46)	5.8% (3)	
14	13.4% (66)	4.8% (1)		12.9% (62)	16.1% (5)		11.9% (55)	23.1% (12)	
>15	7.9% (39)	14.3% (3)		7.7% (37)	16.1% (5)		7.2% (33)	17.3% (9)	
Maturity status									
Pre-PHV	50.2% (224)	40% (8)	$\chi^2_{(2)}=2.655$ $p=0.265$	49.8% (218)	50% (14)	$\chi^2_{(2)}=8.642$ $p=0.06$	52.3% (219)	27.7% (13)	$\chi^2_{(2)}=17.368$ $p<0.001$
Circa-PHV	29.8% (133)	25% (5)		30.8% (135)	10.7% (3)		29.6% (124)	29.8% (14)	
Post-PHV	20% (89)	35% (7)		19.4% (85)	39.3% (11)		18.1% (76)	42.6% (20)	
BMI classification									
Underweight/Normal	57.9% (285)	42.9% (9)	$\chi^2_{(1)}=1.869$ $p=0.172$	57.5% (277)	54.8% (17)	$\chi^2_{(1)}=0.082$ $p=0.774$	58.8% (271)	44.2% (23)	$\chi^2_{(1)}=4.046$ $p=0.04$
Overweight/Obese	42.1% (207)	57.1% (12)		42.5% (205)	45.2% (14)		41.2% (190)	55.8% (29)	

Note: M: male, F: female, CA: chronological age, PHV: peak height velocity, BMI: body mass index

According to the sex of the participants, girls showed a higher prevalence of BP (NP: 8.2%, MBP: 16.4% and LBP: 21.5%) for NP (girls: 4.7%, boys: 3.5%) and LBP (girls: 12.9%, boys: 7.4%) than boys, finding a significant association between sex and LBP. The crude OR from the binary logistic regression confirmed the results from the chi-square for LBP and female sex (OR: 1.85, 95% CI: 1.02-3.35).

For CA groups, the prevalence in the different areas of the spine increased with age (NP: from 2.6% to 7.1%, MBP: from 5.3% to 11.9%, LBP: from 5.3% to 21.4%), finding a significant association between CA and LBP. Regarding the maturity status, there was also an increase in the prevalence of pain according to the pubertal development of the participants, with the highest increase in the post-PHV state. A significant association was found between being in a post-PHV state and having LBP. The crude OR from the binary logistic regression confirmed the results from the chi-square for LBP and 12- (OR: 3.60, 95% CI: 1.10-11.74), 13- (OR: 3.93, 95% CI: 1.20-12.84), 15-year-old (OR: 4.91, 95% CI: 1.41-17.10) and post-PHV (OR: 4.43, 95% CI: 2.10-9.34).

According to BMI classification, the prevalence of LBP in overweight or obese participants (13.2%) was the only one that increased considerably to their normal-weight peers (7.8%), finding an association between being overweight or obese and presenting LBP (OR: 1.79, 95% CI: 1.01-3.20).

When analysing the possible associations between having pain in one area or another of the spine, an association was found between those with NP and those with MBP ($\chi^2_{(1)}$=7.524, p=0.006, OR: 5.60, 95% CI: 1.90-16.48), and between those with MBP and those with LBP ($\chi^2_{(1)}$=4.406, p=0.03, OR: 2.83, 95% CI: 1.16-6.94).

The prevalence of BP in the last month according to the area of spinal pain and adjusted by sex showed only an association between LBP and CA groups ($\chi^2_{(6)}$=18.521, p=0.003) and maturation stage in girls ($\chi^2_{(2)}$=14.777, p=0.001), in this case, the regression showed an association with 14- (OR: 4.24, 95% CI: 1.01-17.73), 15-year-old (OR: 6.22, 95% CI: 1.45-26.74) and post-PHV (OR: 7.03, 95% CI: 2.24-22.03).

3.3.4. Characteristics of back pain

Nearly half of participants (48.9%) reported having BP “a few days” in the past year. Of the 180 students with BP, only 22 (12.3%) experienced pain that prevented them from performing activities of daily living (ADL) and 11 (6.1%) indicated suffering sciatica. As for the positions in which they experienced BP more frequently, they were the following: standing (38.3%), when bending the trunk forward (31.7%), seated (30%) and lying down (8.3%). The mean intensity of BP reported was 3.87±2.01 on the VAS. Non-recurrent (at least once) pain was experienced on average as mild (2.7±1.7 VAS points), occasional (a few days) (4.3±1.8 VAS points) and frequent (often) pain (5.6±1.4 VAS points) as moderate, while daily pain was severe (6.3±2.1 VAS points).

There was an association between those who reported ADL limitation and the frequency of pain ($\chi^2_{(3)}$=19.223, p<0.001). Students who presented BP during the past year and reported frequencies of "a few days" or "daily" presented an association with the limitation of ADL. In addition, adolescents who reported ADL limitation due to their spinal pain indicated significantly higher pain intensity (5.1±2.1 VAS points) than those whose spinal pain did not affect daily life (3.7±1.9 VAS points) (p = 0.001). Differences were also found between the frequency and intensity of pain, being greater the intensity of BP in those who reported more frequency of episodes ($F_{(3,178)}$=20.38, p<0.001).

No associations were found between the characteristics of BP and sex, nor were differences found between pain intensity and CA groups, maturation status or BMI classification. However, there were associations between frequency and CA groups ($\chi^2_{(18)}$=31.682, p=0.02), maturational status ($\chi^2_{(6)}$=12.397, p=0.03) and BMI classification ($\chi^2_{(3)}$=8.878, p=0.02). In particular, there were associations between 13-, 14-year-old, post-PHV and overweight or obese students with frequent pain (often). On the other hand, there was an association between ADL limitation and overweight/obese status ($\chi^2_{(1)}$=13.257, p<0.001).

Table 10 presents data on the characteristics of BP in the last year according to the spinal pain area. No association were found between the different spinal pain areas and the frequency, except for the NP and suffering pain "often" ($\chi^2_{(1)}$=7.825, p=0.04, OR: 4.35, 95% CI: 1.45-12.99). The students with pain in more than one spinal area indicated a higher proportion of occasional (64%) than those with MBP (55.2%), LBP (52.8%), or NP (44.7%). However, LBP led to more limitations in daily life (LBP: 17.9%, pain in more than one spinal area: 16%, NP: 10.8%, MBP: 6%), finding an association between having LBP and limitation ($\chi^2_{(1)}$=7.374, p=0.006, OR: 5.04, 95% CI: 1.43-17.71). The intensity of LBP (3.67±2.18 VAS points) and neck (4.03±1.65 VAS points) was the lowest on average, followed by pain in multiple spinal areas (4.25±1.75 VAS points) and MBP (4.26±1.75 VAS points). The one-way ANOVA did not reveal a significant difference in pain intensity for the different regions ($F_{(3,178)}$=1.75, p=0.158).

Table 10. Characteristics of BP in the last year according to the spinal pain area.

		Neck pain (38)	Mid-back pain (67)	Low back pain (106)	Pain in more than one spinal area (25)
Frequency	At least once	31.6% (12)	34.3% (23)	36.8% (39)	20% (5)
	A few days	44.7% (17)	55.2% (37)	52.8% (56)	64% (16)
	Often	23.7% (9)	9% (6)	8.5% (9)	16% (4)
	Daily	0% (0)	1.5% (1)	1.9% (2)	0% (0)
BP positions	Lying down	15.8% (6)	9% (6)	8.4% (9)	16% (4)
	Seated	36.8% (14)	31.3% (21)	28% (30)	36% (9)
	Standing	36.8% (14)	37.3% (25)	43% (46)	48% (12)
	When bending trunk forward	26.3% (10)	35.8% (24)	32.7% (35)	28% (7)
ADL Limitation	No	89.2% (34)	94% (63)	82.1% (87)	84% (21)
	Yes	10.8% (4)	6% (4)	17.9% (19)	16% (4)
Sciatica	No	92.1% (35)	92.5% (62)	92.5% (98)	84% (21)
	Yes	7.9% (3)	7.5% (5)	7.5% (8)	16% (4)
VAS	mean±SD	4.03±1.65	4.26±1.75	3.67±2.18	4.25±1.75

Note: BP: back pain, ADL: activities of daily living, VAS: visual analogue scale.

3.4. Discussion

This study aimed to describe the prevalence of BP and its characteristics in a sample of children and adolescents from the Region of Murcia and explore the possible associations between participant's characteristics and BP.

3.4.1. Back pain prevalence in the last year and last month

More than a third of the participants (35.1%) indicated having had BP during the previous year, finding a similar prevalence among boys (45%) and girls (55%). BP was greater among the older (post-PHV) (OR: 2.73, 95% CI: 1.67-4.47) and overweight or obese (OR: 1.58, 95% CI: 1.09-2.28) participants. If CA is taken into account, two ages seem critical since associations were found between BP and 12-year-old (OR: 1.18, 95% CI: 1.18-4.64) and 15-year-old (OR: 3.49, 95% CI: 1.58-7.7) participants. With the prevalence of BP during the last month, the same trend appeared for CA and maturity, although with a lower prevalence (17%), but not with BMI classification. However, an association was found between girls and BP (OR: 1.6, 95% CI: 1.01-2.55).

The 1-year and 1-month mean prevalence of BP found in the present study show lower values concerning those found in other studies. For example, Kedra et al. (2013) reported a mean BP prevalence of 76.2% in Polish students during the preceding year (10-19 years) [36]. Aparicio-Sarmiento et al. (2019) showed a mean 1-year BP prevalence of 55.1% in Spanish high school students (12-17 years), although in Spanish elementary students (8-

12 years) the mean 1-year prevalence was lower (22.3%) [8]. For the last month prevalence, in a Swiss sample, aged from 10 to 16 years old, the mean prevalence of BP was 44.4% [37] and in 9-year-old Danish children, the mean prevalence was 33% [38]. These differences in the mean prevalence between the studies could be due to the design and methodological quality of studies, to discrepancies in the definition of BP, the delimitation of pain areas and the different tools used to diagnose BP, to the disparity of recall periods used, as well as the range of age and characteristics of scholars [1,3].

When the 1-year prevalence of BP was analysed adjusted by sex, the results revealed different patterns. On the one hand, BP in boys was associated with greater BMI (overweight/obesity) (OR: 1.93, 95% CI: 1.14-3.30) and being 12 years old (OR: 2.58, 95% CI: 1.01-6.62), while in girls it was associated with complete pubertal development (post-PHV) (OR: 3.79, 95% CI: 1.93-7.41) and a CA of 13 (OR: 6.07, 95% CI:1.81-20.28), 14 (OR: 2.80, 95% CI: 1.81-7.34) and 15 (OR: 5.25, 95% CI: 1.71-16.11) years. Regarding the prevalence of BP in the last month adjusted by sex, girls presented the same pattern as during the previous year (post-PHV and 13, 14 years). However, boys presented the only association with 12-year-old. The results of the present study partly confirm the hypothesis previously stated. However, when analysing the prevalence adjusted by sex, the hypothesis about the increase in prevalence as the maturational state increases occurs among girls, and the increase in prevalence as weight increases of schoolchildren occurs among boys.

Sex-related differences in the prevalence of BP have always been present with most studies finding a higher risk of BP in females [39]. Taking into account that BP in girls was found associated with CA (from 13 years) and maturity status (post-PHV), it could be said that back troubles could develop or, at least, become more readily felt during the period of puberty than in earlier childhood, as has been suggested by others [11].

The results of the present study coincide with those found by Wedderkopp et al. (2005) in a study on BP during the last month in girls (8-10 years and 14-16 years), where they found an increase in BP in puberty stages 3 and 4 (Tanner's classification) because in these stages is when accelerated growth occurs and could make the back more susceptible to mechanical injuries [40]. However, it is difficult to compare the results of this study with others because the few studies that have analysed the association between BP and pubertal development have used the classification proposed by Tanner [14,40] or the Pubertal Development Scale [10,17,41]. In our case, we decided to use the regression equation proposed by Mirwald et al. (2002) since it is a non-invasive method to calculate the maturity offset. There are several studies on BP and spinal development that use the maturity offset, to our knowledge, but only with a pre-pubertal sample [42–44].

Another explanation could be since girls mature earlier than boys, and the hormonal changes that appear can alter their pain modulation, contributing to make them more susceptible to pain conditions [26,39,45]. To these aspects, it should be added that during the pubertal development the children's musculoskeletal structures are not mature enough to support the loads derived from growth, and it is also a critical period where most of scoliosis or asymmetries of the body axis appear, which may contribute at the appearance or increase of BP [19,45]. However, there is a need to further investigation about what circumstances during the onset of puberty may induce an increase in reports of BP in girls.

In contrast, no association was found between BP and the maturation stage of the boys, but rather between BP and being overweight or 12 years old. It has to be noted that the onset of puberty in boys is established around 14 years [13], so the association between the pain and 12 years must be due to other factors. In a recent longitudinal study, the association of BP with weight status was studied, among others, during childhood (from 4 years to 15 years). Individuals who were classified as overweight or obese had a higher risk of developing BP at any site. Furthermore, adjusting by age, sex, socioeconomic level and nationality, a child with obesity showed a 34% higher risk of developing BP before the age of 15 [2]. Returning to the results of the present study, a higher percentage of boys was found to be overweight or obese at 12 years of age (55%). The reasons for the increased risk of BP in overweight or obese children are unknown, but the aetiology may be mechanical, biological, or psychological [2,46].

3.4.2. Prevalence of back pain by spinal area in the last year and last month

The spinal area where the BP (1-year and 1-month prevalence) was most often reported was the lower back, followed by the mid-back and neck area. As for general BP, the prevalence in the different areas of the spine increased with age, pubertal development and BMI classification, but these associations only appeared with the LBP.

In other studies, NP and MBP were also relatively low and showed less association with age, and LBP was the most frequently reported location and where the largest increase with age was observed [38,40,42,47,48]. However, other authors found that in young children (8-12 years), MBP is reported more frequently [8,38,40], while LBP is more common from the age of 13.

Palmer et al. (2020) demonstrated that children who are overweight or obese have an increased risk of developing BP in the lumbar, thoracic and cervical spine [2], like in the present study where overweight or obese participants showed an association with having suffered from LBP in the last year (OR: 1.80, 95% CI: 1.17-2.77) and last month (OR: 1.79, 95% CI: 1.01-3.20). As stated above, being overweight or obese could increase the mechanical load on the spine by increasing compressive forces on the immature structures of the spine and, in turn, interfere with the nutrition of the intervertebral disc and increase the risk of LBP [19,26,27].

The prevalence of NP, MBP and LBP was higher among girls. In line with other studies [26,27,48] which reported female sex to be a significant risk factor for LBP (OR from 2.05 to 2.5) and MBP (OR: 4.3, 95% CI: 1.62-11.44), girls resulted from this study with 1.82 more likely to present MBP in the last year than boys (95% CI: 1.07-3.09) and had 1.85 more probability to present LBP in the last month (95% CI: 1.02-3.35). The location of pain adjusted by sex showed an association for LBP with overweight or obese boys in the last year (OR: 1.89, 95% CI: 1.02-3.49) and with older girls (> 13 years or post-PHV) in the last year (OR: 2.59, 95% CI: 1.18-5.65) and last month (OR: 7.03, 95% CI: 2.24-22.03). Previous studies have found similar associations between being a girl and having LBP, some explanations could be the greater lumbar curvature they present compared to boys of the same age [8,49] or the hormonal changes that appear with pubertal development and could influence the perception of pain [26,39,45]. In a recent study [50] it was observed that menarche was a predictive demographic or medical history factor associated with LBP in gymnasts (6 to 18 years old), for which another reason is that girls reported LBP as a premenstrual symptom, although during the explanations of the questionnaire it was insisted not to do so. On the other hand, as a consequence of the roles set by education and society, boys probably avoid showing symptoms or feelings of pain, while in girls it is more accepted [51,52].

In the present study, those who reported NP were associated with having MBP (last year: OR=3.524, 95% CI=1.68-7.38; last month: OR=5.60, 95% CI=1.90-16.48), and between those with MBP and those with LBP in the last month (OR=2.83, 95% CI=1.15-6.94). Although the NP and MBP prevalence was low compared to LBP, they should not be underestimated as having suffered pain in one area has been associated with having pain in another.

3.4.3. Characteristics of back pain

Regarding the characteristics of pain, both sexes behaved similarly (no differences or associations were found between the variables and sex). The frequency with which they suffered BP was occasional (a few days) with a mean intensity of 3.87±2.01 VAS (moderate). A higher frequency of BP episodes was linked to greater intensity and limitation of ADL, and adolescents who reported ADL limitation due to their spinal pain indicated significantly higher pain intensity. In the study of Aartun et al. (2014) found that the mean pain intensity was lower for students who reported pain "once or twice" and that it increased progressively and significantly in the "sometimes" and "often" groups [6]. Similarly, Wirth and Humphreys (2015) showed that around one-quarter of adolescents with BP reported frequent pain of moderate or severe intensity [37]. These same authors indicate that, regardless of the location of pain, the increase in pain intensity and frequency are predictors of whether the experience of BP had an impact on the daily life of adolescents [37].

On the other hand, higher frequencies were observed as the participants were older (post-PHV) or had overweight/obese. The increased frequency of BP as adolescents age or have a more advanced pubertal development is consistent with previous studies [6,14,36–38]. No differences were found between pain intensity and BP areas, but an association was found between LBP and ADL limitation despite presenting the lowest intensity.

Finally, some limitations of the study should be pointed out. First, it is a cross-sectional study that only shows the association between possible risk factors for BP, so it is not possible to demonstrate a cause-effect relationship. Another limitation is the recall period analysed, the use of 1-year prevalence could cause a great loss of information because it depends on the ability of the participants to recall pain. Therefore, for future research, the lifetime prevalence will be used together with the 1-month prevalence or 1-week, as well as for the characteristics of pain that will be asked for the prevalence of a month or a week. On the other hand, when analysing the different spinal regions adjusted by sex for the CA groups, the maturation stage and the BMI classification, the number of participants in each group is reduced and therefore the OR could be inflated. However, since the trend in the results is the same as in the unadjusted spinal area analysis, the association is considered to be true. Furthermore, sex hormone levels or information about the menarche were not measured, and it could influence the relationship between factors. Lastly, the fact that the sample belongs to the same Spanish area does not allow the results to be extrapolated to other areas of Spain or other countries, since they are influenced by environmental, social, cultural and genetic factors specific to each region. However, the results of the study could be used within the community to promote prevention programs among educational centres.

3.5. Conclusion

In summary, the present study showed a prevalence of BP associated with the maturational state and the weight of participants, finding different prevalence patterns after adjusted by sex. BP as well as LBP in boys, was associated with being overweight or obese, while in girls it was associated with greater pubertal development. The characteristics of BP were also associated with weight and maturation status, regardless of sex, finding a higher frequency and limitation of ADLs among older or overweight/obese participants. Therefore, the present study supports the statement that the prevalence of BP should be reported by spinal area, sex and stage of maturation and complemented with its characteristics, frequency and intensity as a minimum, to offer a complete view of the evolution and severity of BP in young population.

Weight is a modifiable risk factor whose optimization might generate benefits on the musculoskeletal health of schoolchildren and, although sex, age and state of maturity are biological and non-modifiable factors, they must all be taken into account at the time to carry out educational prevention programs.

3.6. References

1. Beynon, A.M.; Hebert, J.J.; Lebouef-Yde, C.; Walker, B.F. Potential risk factors and triggers for back pain in children and young adults. A scoping review, part I: Incident and episodic back pain. *Chiropr. Man. Ther.* 2019, *27*(58), 1-7, doi: 10.1186/s12998-019-0280-9.

2. Palmer, A.; Poveda, J.; Martinez-Laguna, D.; Reyes, C.; de Bont, J.; Silman, A.; et al. Childhood overweight and obesity and back pain risk: a cohort study of 466997 children. *BMJ Open.* 2020, *10*, (e036023), doi:10.1136/ bmjopen-2019-036023.

3. Calvo-Muñoz, I.; Gómez-Conesa, A.; Sánchez-Meca, J. Prevalence of low back pain in children and adolescents: A meta-analysis. *BMC Pediatr.* 2013, *13*(14), 1–12, doi: 10.1186/1471-2431-13-14.

4. Scarabottolo, C.; Pinto, R.; Oliveira, C.; Zanuto, E.; Cardoso, J.; Christofaro, D. Back and neck pain prevalence and their association with physical inactivity domains in adolescents. *Eur. Spine J.* 2017, *26*(9), 2274–80, doi: 10.1007/s00586-017-5144-1.

5. Wirth, B.; Potthoff, T.; Rosser, S.; Humphreys, B.K.; De Bruin, E.D. Physical risk

factors for adolescent neck and mid back pain: A systematic review. *Chiropr. Man. Ther.* 2018, *26*(1), 1–10, doi: 10.1186/s12998-018-0206-y.

6. Aartun, E.; Hartvigsen, J.; Wedderkopp, N.; Hestbaek, L. Spinal pain in adolescents: prevalence, incidence, and course: a school-based two-year prospective cohort study in 1,300 Danes aged 11-13. *BMC Musculoskelet. Disord.* 2014, *15*(187), 1-8, doi: 10.1186/1471-2474-15-187.

7. Aparicio-Sarmiento, A.; Rodríguez-Ferrán, O.; Martínez-Romero, M.T.; Cejudo, A.; Santonja, F.; Sainz de Baranda, P. Back pain and knowledge of back care related to physical activity in 12 to 17 year old adolescents from the region of Murcia (Spain): ISQUIOS programme. *Sustain.* 2019. *11*(19), 5249, doi: 10.3390/su11195249.

8. Sainz de Baranda, P.; Andújar, P.; Collazo-Diéguez, M.; Pastor, A.; Santonja-Renedo, F.; Martínez-Romero, M.T; et al. Sagittal standing spinal alignment and back pain in 8 to 12-year-old children from the Region of Murcia, Spain: ISQUIOS programme. *J. Back Musculoskelet. Rehabil.* 2020, *33*(6), 1003-1014, doi: 10.3233/BMR-191727.

9. Meziat Filho, N.; Coutinho, E.S.; Azevedo e Silva, G. Association between home posture habits and low back pain in high school adolescents. *Eur. Spine J.* 2015, *24*(3), 425–33, doi: 10.1007/s00586-014-3571-9.

10. Janssens, K.A.M.; Rosmalen, J.G.M.; Ormel, J.; Verhulst, F.C.; Hunfeld, J.A.M.; Mancl, L.A.; et al. Pubertal status predicts back pain, overtiredness, and dizziness in american and dutch adolescents. *Pediatrics.* 2011, *128*(3), 553–559, doi: 10.1542/peds.2010-2364.

11. Lardon, A.; Leboeuf-Yde, C.; Le Scanff, C.; Wedderkopp, N. Is puberty a risk factor for back pain in the young? A systematic critical literature review. *Chiropr. Man. Ther.* 2014, *22*(1), 1–12, doi: 10.1186/s12998-014-0027-6.

12. Kamper, S.; Henschke, N.; Hestbaek, L.; Dunn, K.; Williams, C. Musculoskeletal pain in children and adolescents. *Brazilian J. Phys. Ther.* 2016, *20*(3), 275–284, doi: 10.1590/bjpt-rbf.2014.0149.

13. Malina, R.M.; Bouchard, C.; Bar-Or, O. *Growth, Maturation, and Physical Activity*, 2nd ed.; Human Kinetics: Champaign, USA, 2004.

14. Hebert, J.J.; Leboeuf-Yde, C.; Franz, C.; Lardon, A.; Hestbæk, L.; Manson, N.; et al. Pubertal development and growth are prospectively associated with spinal pain in young people (CHAMPS study-DK). *Eur. Spine J.* 2019, *28*(7), 1565–1571, doi: 10.1007/s00586-019-05905-6.

15. Garcia, C.; Teles, J.; Barrigas, C.; Fragoso, I. Health-related quality of life of Portuguese children and adolescents according to their biological maturation and volume of physical

activity. *Qual. Life Res.* 2018, *27*(6), 1483–1492, doi: 10.1007/s11136-018-1822-7.

16. Drenowatz, C.; Greier, K. Association of biological maturation with the development of motor competence in Austrian middle school students - A 3-year observational study. *Transl. Pediatr.* 2019, *8*(5), 402–411, doi: 10.21037/TP.2019.09.03.

17. LeResche, L.; Mancl, L.A.; Drangsholt, M.T.; Saunders, K.; Von Korff, M. Relationship of pain and symptoms to pubertal development in adolescents. *Pain.* 2005, *118*(1–2), 201–209, doi: 10.1016/j.pain.2005.08.011.

18. Lloyd, R.S.; Oliver, J.L.; Faigenbaum, A.D.; Myer, G.D.; De Ste Croix, M.B. Chronological age vs. biological maturation: implications for exercise programming in youth. *J. Strength Cond. Res.* 2014, *28*(5), 1454–1464, doi: 10.1519/JSC.0000000000000391.

19. Dos Santos, M.A.; Lunkes, L.C.; Ribeiro, A. de O.; Souza, A. de C. Low back pain and risk factors during the third infancy. *Fisioter. Mov.* 2019, *32*, e003210, doi: 10.1590/1980-5918.032.ao10.

20. John, C.; Rahlf, A.L.; Hamacher, D.; Zech, A. Influence of biological maturity on static and dynamic postural control among male youth soccer players. *Gait Posture.* 2019, *68,* 18–22, doi: 10.1016/j.gaitpost.2018.10.036.

21. MacDonald, J.; Stuart, E.; Rodenberg, R. Musculoskeletal low back pain in school-aged children a review. *JAMA Pediatr.* 2017, *171*(3), 280–287, doi: 10.1001/jamapediatrics.2016.3334.

22. Shultz, S.J.; Nguyen, A.D.; Schmitz, R.J. Differences in lower extremity anatomical and postural characteristics in males and females between maturation groups. *J. Orthop. Sports Phys. Ther.* 2008, *38*(3), 137–149, doi: 10.2519/jospt.2008.2645.

23. Round, J.M.; Jones, D.A.; Honour, J.W.; Nevill, A.M. Hormonal factors in the development of differences in strength between boys and girls during adolescence: A longitudinal study. *Ann. Hum. Biol.* 1999, *26*(1), 49–62, doi: 10.1080/030144699282976.

24. Dejanovic, A.; Cambridge, E.D.J.; McGill, S. Isometric torso muscle endurance profiles in adolescents aged 15-18: normative values for age and gender differences. *Ann Hum Biol.* 2014, *41*(2), 153–158, doi: 10.3109/03014460.2013.837508.

25. Ervin, R.; Fryar, C.; Wang, C-Y.; Miller, I.; Ogden, C. Strength and body weight in US children and adolescents. *Pediatrics.* 2014, *134*(3), e782–789, doi: 10.1542/peds.2014-0794.

26. Minghelli, B.; Oliveira, R.; Nunes, C. Non-specific low back pain in adolescents from

the south of Portugal: prevalence and associated factors. *J Orthop Sci.* 2014, *19*(6), 883–892, doi: 10.1007/s00776-014-0626-z.

27. Graup, S.; de Araújo Bergmann, M.L.; Bergmann, G.G. Prevalence of nonspecific lumbar pain and associated factors among adolescents in Uruguaiana, state of Rio Grande do Sul. *Rev. Bras. Ortop.* 2014, *49*(6), 661–667 doi: 10.1016/j.rboe.2014.10.003.

28. Fogel, B.; Greenberg, D. *Psychiatric Care of the Medical Patient Psychiatric.* In: Care of the Medical Patient. Oxford Uni: Oxford, UK, 2015.

29. Mirwald, R.L.; Baxter-Jones, A.D.G.; Bailey, D.A.; Beunen, G.P. An assessment of maturity from anthropometric measurements. *Med. Sci. Sports Exerc.* 2002, *34*(4), 689–694, doi: 10.1097/00005768-200204000-00020.

30. Radnor, J.M.; Oliver, J.L.; Waugh, C.M.; Myer, G.D.; Lloyd, R.S. The influence of maturity status on muscle architecture in school-aged boys. *Pediatr. Exerc. Sci.* 2020, *32*(2), 89–96, doi: 10.1123/pes.2019-0201.

31. Robles-Palazón, F.J.; Ayala, F.; Cejudo, A.; De Ste Croix, M.; Sainz de Baranda, P.; Santonja, F. Effects of Age and Maturation on Lower Extremity Range of Motion in Male Youth Soccer Players. *J. Strength Cond. Res.* 2020, Publish Ahead of Print, doi: 10.1519/JSC.0000000000003642.

32. Dionne, C.E.; Dunn, K.M.; Croft, P.R.; Nachemson, A.L.; Buchbinder, R.; Walker, B.F.; et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine (Phila Pa 1976).* 2008, *33*(1), 95–103, doi: 10.1097/BRS.0b013e31815e7f94.

33. Sainz de Baranda, P.; Cejudo, A.; Martínez-Romero, M.T.; Aparicio-Sarmiento, A.; Rodríguez-Ferrán, O.; Collazo-Diéguez, M.; et al. Sitting posture, sagittal spinal curvatures and back pain in 8 to 12-year-old children from the region of murcia (Spain): ISQUIOS programme. *Int. J. Environ. Res. Public Health.* 2020, *17*(7), 1–18, doi: 10.3390/ijerph17072578.

34. Martínez-Crespo, G.; Rodríguez-Piñer, M.; López-Salguero, A.I.I.; Zarco-Periñan, M.J.J.; Ibáñez-Campos, T.; Echevarría-Ruiz de Vargas, C.; et al. Dolor de espalda en adolescentes: prevalencia y factores asociados. *Rehabilitación.* 2009, *43*(2), 72–80, doi: 10.1016/S0048-7120(09)70773-X.

35. Rodríguez-Oviedo, P.; Ruano-Ravina, A.; Pérez-Ríos, M.; García, F.B.; Gómez-Fernández, D.; Fernández-Alonso, A.; et al. School children's backpacks, back pain and back pathologies. *Arch. Dis. Child.* 2012, *97*(8), 730–732, doi: 10.1136/archdischild-2011-301253.

36. Kędra, A.; Czaprowski, D. Epidemiology of back pain in children and youth aged 10-19 from the area of the southeast of Poland. *Biomed. Res. Int.* 2013, *2013*, 506823, doi: 10.1155/2013/506823.

37. Wirth, B.; Humphreys, B.K. Pain characteristics of adolescent spinal pain. *BMC Pediatr.* 2015, *15*(42), 1-11, doi: 10.1186/s12887-015-0344-5.

38. Kjaer, P.; Wedderkopp, N.; Korsholm, L.; Leboeuf-Yde, C. Prevalence and tracking of back pain from childhood to adolescence. *BMC Musculoskelet. Disord.* 2011, *12*(98), 1-11, doi: 10.1186/1471-2474-12-98.

39. Beynon, A.M.; Hebert, J.J.; Lebouef-Yde, C.; Walker, B.F. Potential risk factors and triggers for back pain in children and young adults. A scoping review, part II: unclear or mixed types of back pain. *Chiropr. Man. Ther.* 2019, *27*(61), 1–12, doi: 10.1186/s12998-019-0281-8.

40. Wedderkopp, N.; Andersen, L.B.; Froberg, K.; Leboeuf-Yde, C. Back pain reporting in young girls appears to be puberty-related. *BMC Musculoskelet. Disord.* 2005, *6*(52), 1-5, doi: 10.1186/1471-2474-6-52.

41. Hulsegge, G.; van Oostrom, S.H.; Picavet, H.S.J.; Twisk, J.W.R.; Postma, D.S.; Kerkhof, M.; et al. Musculoskeletal Complaints Among 11-Year-Old Children and Associated Factors. *Am. J. Epidemiol.* 2011, *174*(8), 877–884, doi: 10.1093/aje/kwr205.

42. Dolphens, M.; Cagnie, B.; Coorevits, P.; Vanderstraeten, G.; Cardon, G.; D'Hooge, R.; et al. Sagittal standing posture and its association with spinal pain: A school-based epidemiological study of 1196 flemish adolescents before age at peak height velocity. *Spine (Phila Pa 1976).* 2012, *37*(19), 1657–1666, doi: 10.1097/BRS.0b013e3182408053.

43. Dolphens, M.; Cagnie, B.; Coorevits, P.; Vleeming, A.; Danneels, L. Classification system of the normal variation in sagittal standing plane alignment: A study among young adolescent boys. *Spine (Phila Pa 1976).* 2013, *38*(16), 1003–1012, doi: 10.1097/BRS.0b013e318280cc4e.

44. Dolphens, M.; Cagnie, B.; Coorevits, P.; Vleeming, A.; Vanderstraeten, G.; Danneels, L. Classification system of the sagittal standing alignment in young adolescent girls. *Eur. Spine J.* 2014, *23*(1), 216–225, doi: 10.1007/s00586-013-2952-9.

45. Cavalcante Filho, D.E.A.; de Lima Viana, C.N.; de Santana Cabral, M.P.S.; Cunha, F.V.M.; de Sousa Pacheco, F.; da Silva Brito, A.K.; et al. Low back pain in adolescents: A school screening. *J. Hum. Growth Dev.* 2014, *24*(3), 347–353, doi: 10.7322/jhdg.88975.

46. Hartvigsen, J.; Hancock, M.J.; Kongsted, A.; Louw, Q.; Ferreira, M.L.; Genevay, S.; et al. What low back pain is and why we need to pay attention. *Lancet.* 2018, *391*(10137), 2356–2367, doi: 10.1016/S0140-6736(18)30480-X.

47. González-Gálvez, N.; Vaquero-Cristobal, R.; López-Vivancos, A.; Albaladejo-Saura, M.; Marcos-Pardo, P.J. Back Pain Related with Age, Anthropometric Variables, Sagittal Spinal Curvatures, Hamstring Extensibility, Physical Activity and Health Related Quality of Life in Male and Female High School Students. *Int J Environ Res Public Health.* 2020, *17*(7293), 1–16, doi: 10.3390/ijerph17197293.

48. Wirth, B.; Knecht, C.; Humphreys, K. Spine day 2012: Spinal pain in Swiss school children- epidemiology and risk factors. *BMC Pediatr* . 2013, *13*(159), 1-10, doi: 10.1186/1471-2431-13-159.

49. Feng, Q.; Jiang, C.; Zhou, Y.; Huang, Y.; Zhang, M. Relationship between spinal morphology and function and adolescent non-specific back pain: A cross-sectional study. *J. Back Musculoskelet. Rehabil.* 2017, *30*(3), 625–633, doi: 10.3233/BMR-160544.

50. Sweeney, E.A.; Daoud, A.K.; Potter, M.N.; Ritchie, L.; Howell, D.R. Association between Flexibility and Low Back Pain in Female Adolescent Gymnasts. *Clin. J. Sport Med.* 2019, *29*(5), 379–383, doi: 10.1097/JSM.0000000000000660

51. Diepenmaat, A.C.M.; Van Der Wal, M.F.; De Vet, H.C.W.; Hirasing, R.A. Neck/shoulder, low back, and arm pain in relation to computer use, physical activity, stress, and depression among Dutch adolescents. *Pediatrics.* 2006, *117*(2), 412–416, doi: 10.1542/peds.2004-2766.

52. Jones, G.T.; Macfarlane, G.J. Epidemiology of low back pain in children and adolescents. *Arch. Dis. Child.* 2005, *90*(3), 312–316, doi: 10.1136/adc.2004.056812.

ESTUDIO 2. Physical factors associated with back pain in children and adolescents from the Region of Murcia (Spain): ISQUIOS Program

3.1. Introduction

Back pain (BP) is a very frequent reason for consultation among adults, although it is increasingly common in children, especially older children and adolescents, as various epidemiological studies have brought out that show life prevalence ranging from 1% at 7 years, up to 12-40% at 12 years, which increase and almost double from 12 to 15 years (39-71%), coinciding with the onset of puberty. At the end of adolescence, the prevalence of BP approaches that of adults [1–3]. These wide and varied ranges of prevalence that appear in the literature are due to the types of prevalence analysed, the different definitions and anatomical delimitations of BP used, the method used to collect the information and the difference in the age groups studied [3–5].

BP can appear in different areas, one of the most common and studied is the lower back, commonly known as low back pain (LBP) [4], defined as "pain that appears between the lower part of the 12th rib and the gluteal folds that is strong enough to limit usual activities or change the daily routine for more than 1 day, may or may not be irradiated to the leg, and does not include the pain of a febrile illness or menstruation" [6,7]. Although the presence of pain in other areas of the back has been less studied, evidence suggests that cervical pain, defined as pain in the back of the neck, and pain in the thoracic area (mid-back), pain in the middle area of the back between T1-T12, they should also be analysed during childhood and adolescence because its incidence and prevalence are also high [6,8,9]. In adolescents between 15 and 19 years old, LBP and neck pain are in the top ten in years lived with disability worldwide [10]. At 9 years of age, the prevalence of thoracic pain exceeds the LBP prevalence, equalling at 15 years [11,12].

Given that not only LBP but also neck and mid-back pain are common in childhood and adolescence, with an increase in their prevalence with age, it is likely that in these years they are more vulnerable to the appearance of a first episode de BP [3,6,10,13,14]. Furthermore, numerous longitudinal studies agree that the presence of BP in adolescence is a predictor for suffering from this condition in adulthood [4,15–18]. Although the aetiology of BP is diverse and multifactorial, numerous studies have attempted to investigate a wide variety of risk factors (personal, psychological, biomechanical, and genetic) for BP in children and adolescents. It is essential to determine those factors that can produce or trigger BP in young people to design effective prevention programs and identify which subjects are in a situation of greater vulnerability for suffering from this condition [6,19,20].

Taking into account the existing literature together with the latest systematic reviews and meta-analysis carried out in this regard, the most probable risk factors or triggers of BP in young people are predominantly biological and non-modifiable (for example, sex, age, pubertal status, family history, height), making them ineligible targets for preventive interventions [21]. On the other hand, there are modifiable and bidirectional factors (for example, weight, BMI, muscular endurance and flexibility, posture, physical activity, sedentary lifestyle, quality and quantity of sleep, smoking, among others), which should be the focus of care for the implementation of preventive programs [21].

There is a broad consensus in the literature that psychosocial factors are associated with LBP in adolescents [2,22–24]. On the other hand, the role of mechanical factors was reported to be less associated, although many studies were only based on body weight, backpack weight, or physical activity [25–27]. However, there is some evidence that physical risk factors may have been underestimated [6,19,28]. The systematic review by Potthoff et al. (2018) shown that there is conclusive evidence that some measurable physical parameters could put adolescents at risk for LBP. This review found correlations between LBP and trunk muscle endurance and, to a lesser extent, with sagittal spinal alignment, spinal mobility, and neurodynamics. The authors recommended that future studies focus on trunk muscle endurance, preferably in combination with sagittal spinal alignment, spinal mobility and neurodynamics as possible modifiable risk factors for the development of LBP in children and adolescence [19]. Therefore, the aims of the present study were a) to study the anthropometric characteristics and “Postural Fitness” (sagittal spinal alignment, mobility of the spine, and range of motion (ROM) of the hip and trunk muscle endurance) in children and adolescents with and without BP according to sex, and b) to determine if these physical factors are possible triggers of BP. The hypotheses were the following: a) the results from “Postural Fitness” protocol will be different between boys and girls, b) the participants with BP will present a worse spinal sagittal alignment, with a reduced ROM and poorer trunk muscle endurance, and c) the physical triggers for BP will differ based on gender.

3.2. Method

3.2.1. Design

It was a cross-sectional study. All measurements from the “Postural Fitness” protocol (Figure 4) were collected before participate in a postural and physical fitness program called

"ISQUIOS Program". The study was conducted during the first term of the 2017–2018 and 2018-2019 school years.

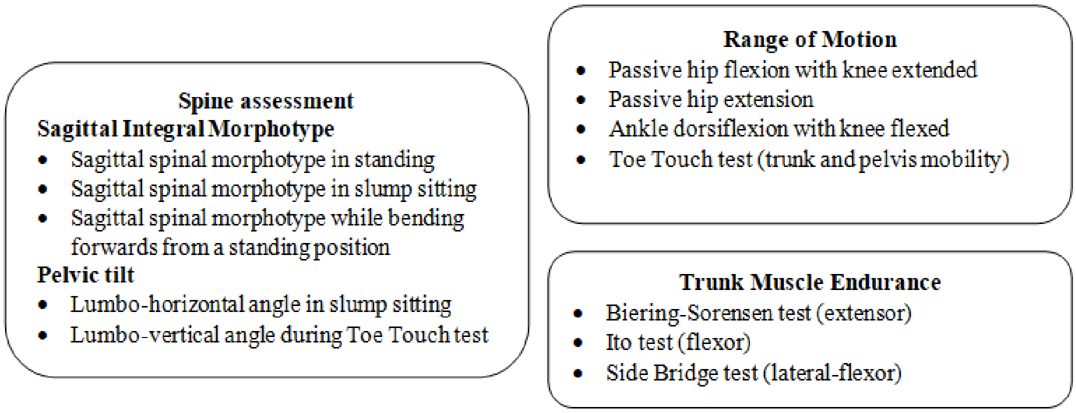


Figure 4. The biomechanical components that belong to the "Postural Fitness" protocol

3.2.2. Participants

A total of 521 students (from 8 to 16 years old) were initially invited from 9 different schools of the Region of Murcia (Spain) to participate in this study (convenience sample). The exclusion criteria were: a) to have a diagnosed spine pathology or important physical injury which limited the correct performance of the tests, b) not to return signed the informed consent (both from parents/guardians and students) before the start of the study, c) to miss the testing session during the data collection, or d) not to have the full "Postural Fitness" assessment.

A comprehensive verbal description of the nature and purpose of the study and the experimental risks was given to the students and their parents/guardians and PE teachers. The protocol was fully approved by the Review Committee for Research Involving Human Subjects at the University of Murcia (Spain) (ID: 1920/2018) and according to the Declaration of Helsinki.

Finally, a sample of 252 (age 10.9±1.2 years (mean±SD), range 9-15 years, 51.6% females) school and high school students completed this study. Three hundred and one students were removed based on the exclusion criteria (Figure 5).

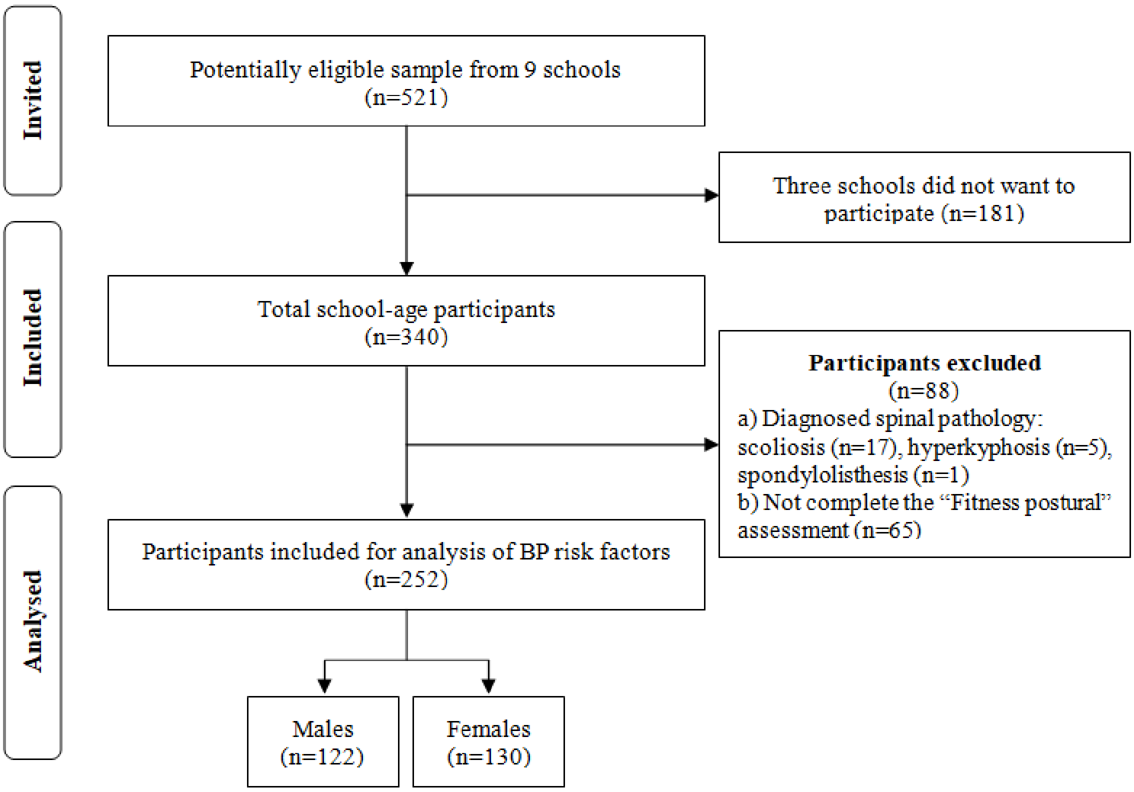


Figure 5. Flow diagram for the sample selection

3.2.3. Procedure

Participants were tested during two different sessions within PE classes. Since PE teachers only have 2 sessions of 60 minutes per age-based grade per week, in the first session (S1) an efficient-time circuit was designed to perform the "Postural Fitness" protocol and anthropometric measures (six different stations were set) (Figure 6). To avoid the appearance of muscular fatigue during the trunk muscle endurance tests, the stations for said tests were alternated with the spine assessment station, the ROM station and the anthropometric station. In this way, between each test for trunk muscles, participants had an average of 5 minutes to rest. The participants were examined wearing sports clothes and barefoot, except for the assessment of the spinal curves that they were in underwear. At the start of the S1, all the participants received comprehensive instructions for the tests, and their questions regarding the protocols were answered. Seven researchers, who were Sports Science specialists with more than 5 years of experience in neuromuscular performance assessments, collected the data at each station on the circuit. Participants were asked not to perform strenuous exercises in the 24h before the testing session.

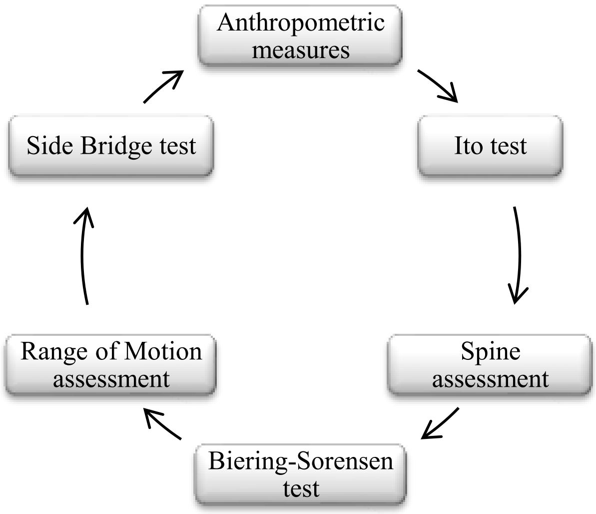


Figure 6. Assessment circuit for the first testing session

In the second session (S2), an experienced researcher presented the BP questionnaire to the parents of each student, explained the procedure to complete the survey and personally solved all participant questions. Students filled out the BP questionnaire with their parents at home. The use of parental reports is very important with school-age children because it helps to improve the quality of the information gathered and has been recommended in prior studies [29].

Anthropometric measures

Body mass in kilograms was measured on a calibrated physician scale (SECA 799, Hamburg, Germany). Body height was recorded in centimetres on a measurement platform (SECA 799). Sitting height was measured in centimetres. Leg length was calculated as the difference between body height and sitting height. Body mass index (BMI) was obtained by dividing body mass by height (in meters) squared and was classified according to the US Centres for Disease Control and Prevention’s age- and sex-matched percentile grading: underweight (equal to or below 5%), normal weight (from 5.01% to 84.99%), overweight (from 85% to 94.99%), and obese (equal to or above 95%). All these measurements were carried out by the same rater within the same station.

Stage of maturation was calculated in a non-invasive manner using a regression equation comprising measures of age, body mass, body height, sitting height and leg length taken during the first part of the testing sessions [30]. Using this method, maturity offset (calculation of years from PHV) was completed (equation 1 for boys and equation 2 for girls). Due to the error in the prediction equation of approximately 6 months in the paediatric population [30], participants with a maturity offset of -0.99 to -0.51 years and +0.51 to +0.99 years were removed from the analysis by maturational stage [31,32]. Therefore, the following equations to calculate maturity offset were used:

Equation 1 (boys) $= -9.236 + 0.0002708 \times (Leg\ Length \times Sitting\ Height) - 0.001663 \times (Age \times Leg\ Length) + 0.007216 \times (Age \times Sitting\ Height) + 0.02292 \times (Weight/Height \times 100)$

Equation 2 (girls) $= -9.376 + 0.0001882 \times (Leg\ Length \times Sitting\ Height) + 0.0022 \times (Age \times Leg\ Length) + 0.005841 \times (Age \times Sitting\ Height) - 0.002658 \times (Age \times Weight) + 0.07693 \times (Weight/Height \times 100)$

"Postural Fitness" protocol

The "Postural Fitness" protocol (Figure 4) is based on the measurement of the biomechanical components that influence on the health of the back through the assessment of the sagittal spinal curvatures and pelvic tilt [33], the hip range of motion (ROM) [34] and the trunk muscles endurance [35–37].

Sagittal Integral Morphotype

Sagittal spinal curvatures (thoracic and lumbar) were examined through the "Sagittal Integral Morphotype" protocol described by Santonja-Medina et al. (2020) to provide a complete assessment of the sagittal spinal alignment (Figure 7). The positions in which the "Sagittal Integral Morphotype" was determined were the following: relaxed standing position (SP), slump sitting position (SSP) and while bending forward from a standing position (BFP). This protocol was created to achieve an accurate and complete diagnosis of spinal sagittal alignment [38–41].

A: Standing position B: Slump Sitting Position C: Bending Forward Position

Figure 7. "Sagittal Integral Morphotype" assessment. A: standing position, B: slump sitting position, C: bending forward position.

A unilevel inclinometer (ISOMED Inc., Portland, OR, USA) was used to quantify the sagittal spinal curvatures. The reference values for each curve and position were in Table 11, negative values stand for degrees of posterior concavity (lordosis), and positive values stand for anterior concavity or kyphosis. Before data collection, the first and twelfth thoracic vertebrae (T1 and T12), and fifth lumbar vertebra (L5) were pointed on the skin of students [33].

Table 11. Reference values for the sagittal spinal assessment of thoracic and lumbar curve in each position

Spinal curve	Standing position		Slump sitting position		Bending forwards from a standing position	
	Values	Classification	Values	Classification	Values	Classification
Thoracic	< 20°	Hypokyphosis	< 20°	Hypokyphosis	< 40°	Hypokyphosis
	20° to 40°	Normal	20° to 40°	Normal	40° to 65°	Normal
	> 40°	Hyperkyphosis	> 40°	Hyperkyphosis	> 65°	Hyperkyphosis
Lumbar	< -20°	Hypolordosis	< -15°	Lordosis	< 10°	Hypokyphosis
	-20° to -40°	Normal	-15° to 15°	Normal	10° to 30°	Normal
	> 40°	Hyperlordosis	> 15°	Hyperkyphosis	> 30°	Hyperkyphosis

<u>Standing Position Assessment (SP)</u>

Participants were placed standing up and relaxing (with eyes and ears aligned horizontally, arms relaxed at the sides of the body, knees extended, and feet hip-width apart) [33]. First, the inclinometer was placed at T1 and calibrated to 0◦, from there the column was profiled until the maximum angulation of the thoracic curvature was reached and the angle was recorded. From this point, the inclinometer was recalibrated to 0◦ and the lumbar spine was outlined until the maximum lumbar angle was reached and recorded (Figure 7A).

<u>Slump Sitting Position Assessment (SSP)</u>

The students were seated on a stretcher with their hands resting on their thighs, and their legs hanging over the stretcher [33]. The thoracic curve was evaluated by placing the inclinometer at T1 and calibrated to 0◦, from there the inclinometer was placed at T12 and the value was recorded. For the lumbar curve, the inclinometer was recalibrated to 0◦ at this last point (T12), the inclinometer was set at L5 and the degrees were recorded (Figure 7B).

<u>Bending Forward Position Assessment (BFP)</u>

The participants were asked to perform the “Toe Touch” test and hold the maximum trunk forward bending for 6-8 seconds with their knees, arms, and fingers fully extended. The students stood on a 36 cm high box with their feet hip-width apart and barefoot. In this position, sagittal spinal curves were assessed following the same procedure as in the SSP [24,46,56,60] (Figure 7C).

Table 12. Diagnostic classification of the "Sagittal Integral Morphotype" for the thoracic curve.

Classification	Subclassification	SP	SSP	BFP
Normal kyphosis		Normal (20° to 40°)	Normal (20° to 40°)	Normal (40° to 65°)
Functional Thoracic Hyperkyphosis	Static	Normal (20° to 40°)	Hyperkyphosis (> 40°)	Normal (40° to 65°)
	Dynamic	Normal (20° to 40°)	Normal (20° to 40°)	Hyperkyphosis (> 65°)
	Total	Normal (20° to 40°)	Hyperkyphosis (> 40°)	Hyperkyphosis (> 65°)
Hyperkyphosis	Total	Hyperkyphosis (> 40°)	Hyperkyphosis (> 40°)	Hyperkyphosis (> 65°)
	Standing	Hyperkyphosis (> 40°)	Normal (20° to 40°)	Normal (40° to 65°)
	Static	Hyperkyphosis (> 40°)	Hyperkyphosis (> 40°)	Normal (40° to 65°)
	Dynamic	Hyperkyphosis (> 40°)	Normal (20° to 40°)	Hyperkyphosis (> 65°)
Hypokyphosis // Hypokyphotic attitude	Flat-back	Hypokyphosis (< 20°)	Hypokyphosis (< 20°)	Hypokyphosis (< 40°)
	Standing	Hypokyphosis (< 20°)	Normal (20° to 40°)	Normal (40° to 65°)
	Static	Hypokyphosis (< 20°)	Hypokyphosis (< 20°)	Normal (40° to 65°)
	Dynamic	Hypokyphosis (< 20°)	Normal (20° to 40°)	Hypokyphosis (< 40°)
Hypomobile kyphosis		Normal (20° to 40°)	Normal (20° to 40°)	Hypokyphosis (< 40°)

Note: SP: Standing position; SSP: Slump sitting position; BFP: Bending forward position.

Tables 12 and 13 show the different classification and subclassification for the comprehensive thoracic and lumbar sagittal diagnosis, respectively. This diagnosis is established from the values of the 3 positions measured.

Table 13. Diagnostic classification for the "Sagittal Integral Morphotype" for the lumbar curve.

Classification	Subclassification	SP	SSP	BFP
Normal lordosis		Normal (-20° to -40°)	Normal (-15° to 15°)	Normal (10° to 30°)
Lumbar spine with reduced mobility	Functional lumbar lordosis // Hypomobile lordosis	Normal (-20° to -40°)	Normal (-15° to 15°)	Hypokyphosis or lordosis (< 10°)
	Lumbar hypomobility	Hypolordosis (< -20°)	Normal (-15° to 15°)	Hypokyphosis (< 10°)
Hyperlordotic attitude		Hyperlordosis (> -40°)	Normal (-15° to 15°)	Normal (10° to 30°)
Functional lumbar hyperkyphosis	Static	Normal (-20° to -40°)	Hyperkyphosis (> 15°)	Normal (10° to 30°)
	Dynamic	Normal (-20° to -40°)	Normal (-15° to 15°)	Hyperkyphosis (> 30°)
	Total	Normal (-20° to -40°)	Hyperkyphosis (> 15°)	Hyperkyphosis (> 30°)
	Hypermobility 1	Hyperlordosis	Hyperkyphosis	Hyperkyphosis

Table 13. Diagnostic classification for the "Sagittal Integral Morphotype" for the lumbar curve.

Classification	Subclassification	SP	SSP	BFP
		(> -40°)	(> 15°)	(> 30°)
Lumbar Hypermobility	Hypermobility 2	Hyperlordosis (> -40°)	Normal (-15° to 15°)	Hyperkyphosis (> 30°)
	Hypermobility 3	Hyperlordosis (> -40°)	Hyperkyphosis (> 15°)	Normal (10° to 30°)
Hypolordosis	Hypolordotic attitude	Hypolordosis (< -20°)	Normal (-15° to 15°)	Normal (10° to 30°)
	Lumbar kyphosis 1	Hypolordosis (< -20°)	Hyperkyphosis (> 15°)	Hyperkyphosis (> 30°)
	Lumbar kyphosis 2	Hypolordosis (< -20°)	Hyperkyphosis (> 15°)	Normal (10° to 30°)
	Lumbar kyphosis 3	Hypolordosis (< -20°)	Normal (-15° to 15°)	Hyperkyphosis (> 30°)
Structured Hyperlordosis		Hyperlordosis (> -40°)	Hyperlordosis (< -15°) or normal (-15° to 15°)	Lordosis or Hypokyphosis (< 10°)
Structured lumbar kyphosis		Hypolordosis or kyphosis (< -20°)	Hyperkyphosis (> 15°)	Hyperkyphosis (> 30°)

Note: SP: Standing position; SSP: Slump sitting position; BFP: Bending forward position.

Pelvic tilt

Lumbo-sacral angle

Pelvic tilt was assessed using the lumbosacral angle (LSA) in a slump sitting position (Figure 8) and during performed the "Toe Touch" test.

Figure 8. Pelvic tilt assessment through Lumbo-Horizontal angle.

Through this test, it could be determined whether the students can maintain the verticality of the pelvis and, consequently, a more neutral sagittal spine when slump sitting or bending forward the trunk. For the assessment, the angle generated by the horizontal (L-H) or vertical (L-V) line and the spinous processes of L4-S1 was measured with a goniometer provided with a spirit level system (Fabrication Enterprises Inc., NY, USA) [42,43]. However, for the data analysis, the supplementary angle was used. Pelvic tilt was classified as normal when the angle was ≤100°, and as posterior pelvic tilt when the angle was >101° [43,44]. Both the spine and pelvic tilt assessment were carried out by the same rater at the same station.

Range of motion

For the evaluation of the ROM, 3 tests from the ROM-SPORT battery were used [34]. Measurements were performed with a unilevel inclinometer (ISOMED Inc., Portland, OR, USA) with an extendable telescopic arm and lumbar protection support (Lumbosant©, Murcia, Spain) was used to maintain the pelvis in a neutral position during the evaluation tests, except for the ankle dorsiflexion with knee flexed test. Two measurements of each ROM test were made for each leg, and the mean score from each test was used in statistical analyses [34]. All these measurements were carried out by the same two raters (principal and assistant) within the same station. The endpoint for each test was determined by 1 or more of the following criteria: a) the rater is unable to continue the stretching movement due to the elevated resistance of the tested muscle groups, b) one or both raters perceive compensatory movements, and c) the participant feels a strong but tolerable stretch, slightly before the occurrence of pain [34].

Passive hip flexion with the knee extended (PHFKE)

The assessment of the PHFKE (or passive straight leg raise test) was carried out with the participants in a supine position on the stretch with their legs extended. The Lumbosant© was placed in the lower back. Before the assessment, the inclinometer was calibrated to 0° with the horizontal. The inclinometer was placed on the external face of the evaluated leg and the telescopic arm was aligned with the bisector of the leg. The main rater was responsible for performing maximum hip flexion slowly and progressively, avoiding knee flexion and external rotation of the leg assessed. The assistant rater was in charge of checking the other leg, that it did not flex the knee or move the pelvis. Once the maximum range of motion was reached, the angle formed by the longitudinal axis of the leg with the horizontal was recorded (Figure 9) [34]. PHFKE was classified as normal when the angle was ≥75°, and as reduced when the angle was lower than <75° [45].

Figure 9. Hip extensor muscles assessment through the passive hip flexion with knee extended

Passive hip extension (PHE)

The PHE (or modified Thomas test) assessment was carried out with the participants in a supine position with buttocks sticking out over the edge of the stretch. To guarantee this position, the students supported the sacrum/coccyx on the edge of the stretch and from there they laid down. The Lumbosant© was placed in the lower back. The participant held the unexamined leg towards the chest (flexing the hip) to keep the pelvis and spine in a neutral position in contact with the Lumbosant©, to avoid arching of the back during the testing. Before the assessment, the inclinometer was calibrated to 0° with the horizontal. The inclinometer was placed on the external face of the evaluated leg and the telescopic arm was aligned with the bisector of the leg. The main rater applied pressure to the anterior thigh to achieve maximum hip extension slowly and progressively. The assistant rater was in charge of checking the other leg, keeping the lumbar area in contact with the Lumbosant©. Once the maximum range of motion was reached, the angle formed by the longitudinal axis of the thigh with the horizontal was recorded (Figure 10) [34]. PHE was classified as normal when the angle was ≥14°, and as reduced when the angle was lower than <13° [46].

Figure 10. Hip flexor muscles assessment through the passive hip extension

"Toe Touch" test

The Toe Touch (TT) test was used to measure the mobility of both the whole spine and the pelvis in the overall motion of bending forward, and indirectly the hamstring muscle flexibility [47,48]. Participants were asked to bend forward (with feet hip-width apart on the Sit and Reach box) as much as possible with their knees, arms, and fingers extended over the measuring scale. The vertical distance was recorded in centimetres. This test complements the L-V angle as it provides information on the respective participation of the mobility of the spine and the pelvis when bending forward the trunk [43]. This test was performed at the spine assessment station by the same rater.

Trunk muscle endurance

Finally, three field-based tests were selected to assess isometric trunk extensor (Biering-Sorensen test) [37], flexor (Ito test) [35] and lateral-flexor (Side Bridge right and left test) [36] endurance. These tests were selected to measure the endurance of various areas of the trunk and so obtain a larger understanding of the overall trunk muscles. During the performance of all field-based tests, participants were strongly encouraged verbally to maintain the position as long as possible. The tests' duration was recorded in seconds. All these tests were carried out by a different rater in each one.

Biering-Sorensen test

The Biering-Sorensen (BS) test was performed with the participant in a prone position with the lower body resting on a test bench and the anterior superior iliac spine aligned at the edge of the test bench. The lower body was attached to the test bench by 2 inextensible straps (knees and ankles). The test consisted of holding the upper body in a horizontal position with arms crossed on the chest while keeping the head in a neutral position for as long as possible, until exhaustion, or until participants lost the correct position more than 3 times [37] (Figure 11).

Figure 11. Trunk extensor endurance assessment through Biering-Sorensen test

Ito test

The Ito test was performed with the participants in a supine position with hips and knees flexed 90° and arms interlaced with hands grasping the opposite elbow. The test consisted of performing a trunk flexion (curl-up) until the elbows touched their thighs and holding this position for as long as possible, until exhaustion. The test ended when the scapulae came in contact with the mat. The position of the original test was modified to normalize the range of motion to the participants' characteristics and thus avoid hip and lower back flexion (sit-up) following the description of Juan-Recio et al. [35,49] (Figure 12).

Figure 12. Trunk flexor endurance assessment through Ito test

Side Bridge test

The Side Bridge right and left (SB-R and SB-L) test was performed with the participants in a lateral position on their side with legs extended. The participants were supported on their elbow and feet, the top foot was placed ahead of the lower foot (with 90º elbow flexion and the arms perpendicular to the mat) while bridging their hips off the mat to maintain an aligned body position. The test finished when the subject lost the aligned postural position [36] (Figure 13).

Figure 13. Trunk lateral-flexor endurance assessment through Side Bridge test

BP Assessment

An ad-hoc questionnaire composed of 8 items was used to describe BP prevalence in school-age males and females [4,7,50–52]. The reliability and validity of the questionnaire have been evaluated in previous studies with Spanish children and adolescents [53,54].

The questionnaire starts with sociodemographic data such as sex, age, school, grade level, and pathologies diagnosed. After that, there are questions about BP prevalence within the last 12 months and last 1-month ("During the last year (or last month), have you had back pain?"). BP was defined as aching, pain, or discomfort in some parts of the back that were not related to trauma or menstrual pain. Those who experienced BP within the last year completed the items based on the spinal pain areas (neck, mid-back and/or low back) and the characteristics of pain within the last year (frequency and intensity of pain, irradiation to lower limb (sciatica) and limitations of daily life activities). The questionnaire included a drawing of the back to mark the most affected area (Figure 14). The frequency of spinal pain was classified as "at least once", "some days", "often (several days)" or "daily". The intensity of spinal pain was quantified using a visual analogue scale (VAS) ranging from 0 (no pain) to 10 (intolerable pain).

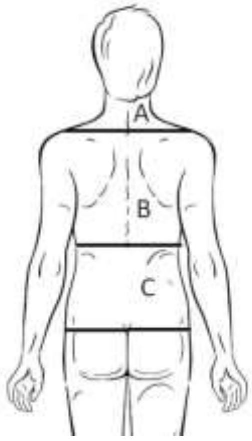


Figure 14. Drawing of the back to mark the area of back pain. A: neck pain, B: mid-back pain, C: low back pain

3.2.4. Statistical Analyses

The normality of the distribution for each variable was determined by the Kolmogorov-Smirnov test. Descriptive analysis was performed for the total sample, as well as by sex, including mean and standard deviations (SD) for the quantitative variables and counts and percentages for qualitative variables.

Independent t-tests (two groups) or ANOVA (more than two groups) were used to explore differences in mean values between independent samples (e.g., males and females) and continuous variables. For non-normally-distributed variables, Mann-Whitney tests (two groups) or Kruskal-Wallis tests (more than two groups) were used. The magnitude of the effect size was classified as previously described by Hopkins et al. (2009) as trivial (< 0.2), small (0.2 to 0.59), moderate (0.6 to 1.19), large (1.20 to 2.00), very large (2.00 to 3.99) or extremely large (> 4.0) [55].

Chi-square tests (bivariate analysis) were applied to explore associations between the presence of BP (dependent variable) and the qualitative variables (or Fisher's exact test). Furthermore, to determine the associations between BP and the variables, crude odds ratio (OR) with 95% confidence interval (95% CI) were calculated.

Forward stepwise binary logistic regression models (probability $p \leq 0.05$, elimination probability $p \leq 0.10$) were used to identify variables associated with BP; odds ratio analysis was used for calculating the simultaneous effects of several predictors instead of relative risk estimates [56]. Small effect (from 1 to 1.25), medium effect (from 1.25 to 2), and large effect (≥ 2) were the three categories used to interpret the effect sizes for the odds ratio (OR) [57].

The statistical analysis was conducted using the statistical package SPSS v.21.0 for Windows (IBM Corp., Armonk, NY, USA).

3.3. Results

This study was based on a sample of 252 students, aged from 9 to 15 years. Descriptive statistics for the total sample and by sex are displayed in Table 14.

Table 14. Descriptive values of anthropometric characteristics and "Fitness Postural" assessment (mean ± SD) per total sample and by sex

Variables	Total sample (n=252)	Male (n=122)	Female (n=130)
	Anthropometric Characteristics		
Age (y)	10.9±1.2	11.1±1.4	10.8±0.9*
Maturity offset	-1.5±1.2	-2.2±1.1	-0.9±0.9*
Body mass (kg)	43.8±11.2	44.5±11.3	43.1±11.2
Body height (cm)	145.9±9.5	147±9.9	144.9±9
BMI (kg/m^2)	20.4±4	20.4±3.9	20.3±4.1
Percentile	69.1±29.5	69.6±29.5	68.6±29.5
	Fitness Postural *Spine Assessment*		
SP thoracic (º)	38.8±9.5	40.8±9.2	36.9±9.3*
SP lumbar (º)	33.7±9.6	31.7±9.5	35.7±9.3*
SSP thoracic (º)	36.1±10.8	38.9±10.8	33.5±10.1*
SSP lumbar (º)	11.1±10.7	13±11.5	9.3±9.7*
L-H angle (º)	102.3±8.6	104.8±8.5	99.8±7.9*
BFP thoracic (º)	48.7±11.5	50.2±10.9	47.2±12*
BFP lumbar (º)	26±8.9	27.4±8.9	24.7±8.7*
L-V in BFP (º)	114.5±14.1	120.6±12.9	108.7±12.7*
	Range of Motion		
PHE-R (º)	17.2±8.4	16.1±8.5	18.2±8.2*
PHE-L (º)	17.3±8.3	16.4±8.2	18.2±8.3‡
PHFKE-R (º)	73.9±12.9	69.3±9.6	78.2±14.1*
PHFKE-L (º)	73.1±12.7	68.7±10.2	77.3±13.4*
TT distance (cm)	-7.5±8.7	-10.2±7.8	-5±8.9*
	Trunk Muscle Endurance		
Ito test (s)	83.7±72	86.3±82.1	81.5±61.9
BS test (s)	129.3±71.1	120.9±73.1	137.3±68.5*
SB-R test (s)	44±25.4	47.6±27.5	40.6±23*
SB-L test (s)	45±26.5	50±29.7	40.5±22.6*

Note: y: years, kg: kilograms, cm: centimetres, BMI: body mass index, º: degrees, SP: standing position, SSP: slump sitting position, L-H: lumbo-horizontal, BFP: bending forward position, L-V: lumbo-vertical, ADKF (R-L): ankle dorsiflexion with knee flexed (right-left), PHE (R-L): passive hip extension (right-left), PHFKE (R-L): passive hip flexion with the knee extended (right-left), TT: "Toe Touch" test, BS: Biering-Sorensen, s: seconds, SB (R-L): Side Bridge (right-left). *Significant differences between boys and girls ($p<0.05$). ‡p value=0.06.

Among the main characteristics, it should be noted that 171 students (67.9%) were in a state before the onset of puberty (104 boys and 67 girls) and 81 participants (32.1%) were in the pubertal stage (18 boys and 63 girls). Concerning anthropometric characteristics, the main sex-related differences of the participants were found in age and maturity offset, finding that girls were younger but presented a higher maturity offset. Regarding the results of the "Fitness Postural" protocol, in the assessment of thoracic spinal curvature boys presented a greater kyphosis in all the assessed positions. Conversely, for the lumbar curve, it was the girls who showed greater lordosis in standing position and lower lumbar kyphosis, together with greater neutral pelvis in slump sitting and in bending forward positions. About ROM, girls showed a greater range in all the tests performed. Finally, in the trunk muscle endurance tests, boys showed greater performance in the lateral-flexor muscles and girls in the extensor musculature.

Table 15 shows that most of the students had normal thoracic kyphosis in standing position (57.1%) and slump sitting position (70.6%). However, in the bending forward position, most of the participants showed thoracic hyperkyphosis (72.2%). When analysing the classifications according to sex, the same pattern was found, but standing and slump sitting hyperkyphosis were associated with boys and normal kyphosis was associated with girls. On the other hand, the diagnosis of the "Sagittal Integral Morphotype" revealed that the predominant morphotype was "dynamic functional hyperkyphosis", followed by "normal" morphotype and "total hyperkyphosis", showing an association between "total hyperkyphosis" morphotype and boys, and between "normal" morphotype and girls. For the lumbar curvature, the most of participants showed normal lumbar curve in the three assessed positions both for the total sample and by sex. Although, associations were found between hyperlordosis and girls in standing position and between lumbar hyperkyphosis and boys in slump sitting position. Regarding the pelvic tilt assessment, most of the students presented posterior pelvic tilt, especially when they bending forward. This posterior pelvic tilt was more common in boys than in girls. Similarly, a majority of participants had a reduced ROM in the hip extensors and flexors, being much more prevalent among boys.

Table 15. Percentages (counts) of participants for each spinal curve by assessment position, "Sagittal Integral Morphotype" classification, pelvic tilt classification and prevalence of back pain by total sample and by sex

	Classification	Total sample (n=252)	Male (n=122)	Female (n=130)	Chi-square
		Thoracic curvature			
Standing position	Hypokyphosis	0.8% (2)	0% (0)	1.5% (2)	$\chi^2_{(2)}$6.38 p=0.04
	Normal	57.1% (144)	50.8% (62)	63.1% (82)	
	Hyperkyphosis	42.1% (106)	49.2% (60)	35.4% (46)	
Slump Sitting position	Hypokyphosis	0.8% (2)	0% (0)	1.5% (2)	$\chi^2_{(2)}$=9.09 p=0.005
	Normal	70.6% (178)	63.1% (77)	77.7% (101)	
	Hyperkyphosis	28.6% (72)	36.9% (45)	20.8% (27)	
Bending forward position	Hypokyphosis	0% (0)	0% (0)	0% (0)	$\chi^2_{(2)}$=1.89 p=0.17
	Normal	27.8% (70)	23.8% (29)	31.5% (41)	
	Hyperkyphosis	72.2% (182)	76.2% (93)	68.5% (89)	

Table 15. Percentages (counts) of participants for each spinal curve by assessment position, "Sagittal Integral Morphotype" classification, pelvic tilt classification and prevalence of back pain by total sample and by sex

	Classification	Total sample (n=252)	Male (n=122)	Female (n=130)	Chi-square
Sagittal Integral Morphotype	Normal kyphosis	17.1% (43)	11.5% (14)	22.3% (29)	$\chi^2_{(8)}$=16.16 p=0.03
	Functional Static Hyperkyphosis	1.6% (4)	0.8% (1)	2.3% (3)	
	Functional Dynamic Hyperkyphosis	30.6% (77)	27.9% (34)	33.1% (43)	
	Total Functional Hyperkyphosis	7.9% (20)	10.7% (13)	5.4% (7)	
	Hyperkyphosis standing	6.3% (16)	8.2% (10)	4.6% (6)	
	Hyperkyphosis static	2% (5)	3.3% (4)	0.8% (1)	
	Hyperkyphosis dynamic	16.7% (42)	15.6% (19)	17.7% (23)	
	Total Hyperkyphosis	17.1% (43)	22.1% (27)	12.3% (16)	
	Hypokyphosis	0.8% (2)	0% (0)	1.5% (2)	
	Lumbar curvature				
Standing position	Hypolordosis	6.7% (17)	9.8% (12)	3.8% (5)	$\chi^2_{(2)}$=5.97 p=0.05
	Normal	73% (184)	74.6% (91)	71.5% (93)	
	Hyperlordosis	20.2% (51)	15.6% (19)	24.6% (32)	
Slump Sitting position	Lordosis	0% (0)	0% (0)	0% (0)	$\chi^2_{(2)}$=7.84 p=0.005
	Normal	64.9% (163)	56.2% (68)	73.1% (95)	
	Hyperkyphosis	35.1% (88)	43.8% (53)	26.9% (35)	
Bending forward position	Hypokyphosis	2.8% (7)	3.3% (4)	2.3% (3)	$\chi^2_{(2)}$=3.48 p=0.17
	Normal	71.8% (181)	66.4% (81)	76.9% (100)	
	Hyperkyphosis	25.4% (64)	30.3% (37)	20.8% (27)	
Sagittal Integral Morphotype	Normal lordosis	36.9% (93)	32.8% (40)	40.8% (53)	$\chi^2_{(11)}$=14.26 p=0.18
	Functional lumbar lordosis	2% (5)	2.5% (3)	1.5% (2)	
	Hyperlordotic attitude	17.1% (43)	11.5% (14)	22.3% (29)	
	Functional static hyperkyphosis	12.3% (31)	14.8% (18)	10% (13)	
	Functional dynamic hyperkyphosis	6% (15)	5.7% (7)	6.2% (8)	
	Total Functional hyperkyphosis	15.5% (39)	18% (22)	13.1% (17)	
	Lumbar Hypermobility	2.4% (6)	3.3% (4)	1.5% (2)	
	Hypolordotic attitude	1.6% (4)	1.6% (2)	1.5 (2)	
	Lumbar kyphosis	2.8% (7)	4.1% (5)	1.5% (2)	
	Structured Hyperlordosis	0.8% (2)	0.8% (1)	0.8% (1)	
	Structured lumbar kyphosis	2.4% (6)	4.1% (5)	0.8% (1)	
	Pelvic tilt				
L-H angle	Neutral	44% (111)	31.1% (38)	56.2% (73)	$\chi^2_{(1)}$=15.97 p<0.001
	Posterior PT	56% (141)	68.9% (84)	43.8% (57)	
L-V angle	Neutral	19.8% (50)	7.4% (9)	31.5% (41)	$\chi^2_{(1)}$=23.10 p<0.001
	Posterior PT	80.2% (202)	92.6% (113)	68.5% (89)	
	Range of Motion				
PHE_R	Normal	59.5% (150)	54.1% (66)	64.6% (84)	$\chi^2_{(1)}$=2.89 p=0.08
	Reduced	40.5% (102)	45.9% (56)	35.4% (46)	
PHE_L	Normal	59.5% (150)	53.3% (65)	65.4% (85)	$\chi^2_{(1)}$=3.83

Table 15. Percentages (counts) of participants for each spinal curve by assessment position, "Sagittal Integral Morphotype" classification, pelvic tilt classification and prevalence of back pain by total sample and by sex

	Classification	Total sample (n=252)	Male (n=122)	Female (n=130)	Chi-square
	Reduced	40.5% (102)	46.7% (57)	34.6% (45)	p=0.05
PHF_KE_R	Normal	39.3% (99)	26.2% (32)	51.5% (67)	$\chi^2_{(1)}$=16.90
	Reduced	60.7% (153)	73.8% (90)	48.5% (63)	p<0.001
PHF_KE_L	Normal	44.4% (112)	33.6% (41)	54.6% (71)	$\chi^2_{(1)}$=11.25
	Reduced	55.6% (140)	66.4% (81)	45.4% (59)	p=0.001
		Back pain last year			
Back pain	No	71.8% (181)	71.3% (87)	72.3% (94)	$\chi^2_{(1)}$=0.03
	Yes	28.2% (71)	28.7% (35)	27.7% (36)	p=0.86
Neck pain	No	96.8% (244)	95.9% (117)	97.7% (127)	$\chi^2_{(1)}$=0.66 p=0.49
	Yes	3.2% (8)	4.1% (5)	2.3% (3)	
Mid-back pain	No	89.3% (225)	91% (111)	87.7% (114)	$\chi^2_{(1)}$=0.71
	Yes	10.7% (27)	9% (11)	12.3% (16)	p=0.39
Low back pain	No	81.3% (205)	81.1% (99)	81.5% (106)	$\chi^2_{(1)}$=0.006
	Yes	18.7% (47)	18.9% (23)	18.5% (24)	p=0.93

Note: L-H: lumbo-horizontal, L-V: lumbo-vertical, PT: pelvic tilt, PHF_KE (R-L): passive hip flexion with knee extension (right-left), PHE (R-L): passive hip extension (right-left)

The presence of BP in the last year was observed in 71 (28.2%) students, of students with BP in the last year, 35 were boys (28.7%) and 36 were girls (27.7%). Regarding the spinal area where the pain was most often reported was the lower back (66.2%), followed by the mid-back (38%) and neck area (11.3%), no association was found between BP in any spinal area and sex (Table 17).

Table 16 shows the comparative analysis of the different continuous variables for the total sample of participants with and without BP and according to sex. Significant differences were observed in the variables related to body weight (body mass, BMI and percentile). Participants with BP showed an increase in these variables, finding these differences were also among boys. In the spine assessment, only significant differences were observed between girls with and without BP for lumbar curvature in bending forward position, finding less lumbar kyphosis in girls with BP.

About ROM, significant differences were found between participants with and without BP for PHE. Participants with BP had a lower ROM, these differences were also found in girls and in boys, a trend towards significance was observed (p=0.06). Lastly, in the trunk muscle endurance tests, only significant differences were found between girls with and without BP for the Ito test.

The distribution of participants with and without BP according to the classification of spinal alignment by assessment position is shown in Table 17. Most of the students with and without BP had normal thoracic kyphosis in a standing position (58.6%) and slump sitting position (71.8%). However, in the bending forward position, most of the participants showed thoracic hyperkyphosis (71.8%). When analysing the classifications according to sex, the same pattern was found. For the lumbar curvature, most of the participants showed normal lumbar curve in the three assessed positions both for the total sample and by sex, regardless of BP.

Table 16. Descriptive values of anthropometric characteristics and "Fitness Postural" assessment (mean ± SD) for participants with or without back pain by total sample and sex.

Variables	Total sample (n=252)		Male (n=122)		Female (n=130)	
	No BP (n=181)	BP (n=71)	No BP (n=87)	BP (n=35)	No BP (n=94)	BP (n=36)
Anthropometric Characteristics						
Age (y)	10.9±1.2	11.1±1.3	11.1±1.4	11.2±1.5	10.7±0.9	10.9±1
Maturity offset	-1.6±1.2	-1.4±1.2	-2.3±1.1	-1.9±1.2	-1±0.9	-0.8±1
Body mass (kg)	42.5±10.6	47.1±12.1*	42.9±10.7	48.5±11.8*	42.1±10.6	45.7±12.5
Body height (cm)	145.2±8.9	147.7±10.6	146±9.5	149.4±10.4	144.5±8.3	146±10.6
BMI (kg/m²)	19.9±3.8	21.4±4.2*	19.9±3.7	21.6±4.2*	19.9±3.9	21.2±4.3
Percentile	66.6±30	75.5±27.2*	67±29.6	76.1±28.7	66.2±30.5	75.1±25.9
Postural Fitness						
Spine Assessment						
SP thoracic (º)	38.7±9.2	38.9±10	41±9.3	40.3±9.1	36.6±8.7	37.6±10.8
SP lumbar (º)	33.6±9.3	34±10.3	31.4±9.1	32.3±10.7	35.7±9.1	35.7±9.8
SSP thoracic (º)	36.3±10.7	35.6±11.2	39.4±11.1	37.6±10.2	33.5±9.4	33.7±11.9
SSP lumbar (º)	11.7±11.3	9.5±8.9	14±12.1	10.7±9.6	9.7±10.2	8.4±8.1
L-H angle (º)	102.4±8.8	101.9±7.8	105.2±8.8	103.9±7.7	99.8±8.1	100±7.6
BFP thoracic (º)	48.7±11.9	48.6±10.6	50.6±11.1	49.1±10.5	46.9±12.4	48.1±10.8
BFP lumbar (º)	26.4±9.4	25.1±7.2	27.5±9.6	27.2±7.1	25.4±9.3	23±6.7*
L-V angle (º)	114.3±14.8	115±12.3	120.9±13.4	119.8±11.7	108.1±13.3	110.3±11.1
Range of Motion						
PHE-R (º)	17.9±8.4	15.2±7.9*	16.9±8.7	14.2±7.7	19±8.2	16.1±8.1*
PHE-L (º)	18.1±8.5	15.1±7.3*	17.2±8.6	14.4±6.9	19±8.4	15.8±7.8*
PHFKE-R (º)	74.6±13.6	72±10.9	69.5±10	68.6±8.7	79.3±14.8‡	75.3±11.9
PHFKE-L (º)	73.8±13.5	71.4±10.4	68.9±10.8	68.2±8.7	78.4±14.1‡	74.5±11.1
TT distance (cm)	-7.3±9	-8.1±8	-10.5±8.1	-9.5±7.1	-4.4±8.9‡	-6.7±8.7
Trunk Muscle Endurance						
Ito test (s)	86.3±76.6	77.2±58.5	84.3±87.9	91.5±64.9	88±65.2	64.8±4.9*
BS test (s)	132±70.6	122.5±72.3	124±72.4	113.1±75.3	139.5±68.5	131.5±69.1
SB-R test (s)	45.5±26	39.9±23.5	49.1±30.7	43.9±25.4	42.3±23.5	36.2±21.3
SB-L test (s)	46.5±27.2	41.4±24.5	52.1±60.7	44.9±26.8	41.5±22.8	38.1±22.1

Note: y: years, kg: kilograms, cm: centimetres, BMI: body mass index, º: degrees, SP: standing position, SSP: slump sitting position, L-H: lumbo-horizontal, BFP: bending forward position, L-V: lumbo-vertical, PHE (R-L): passive hip extension (right-left), PHFKE (R-L): passive hip flexion with the knee extended (right-left), TT: "Toe Touch" test, BS: Biering-Sorensen, s: seconds, SB (R-L): Side Bridge (right-left). * Significant differences between those with and without BP ($p<0.05$).

Although, associations were found between BP and the normal lumbar curve in slump sitting position for the total sample and girls and between BP and the normal lumbar curve during bending forward position for the total sample. Regarding pelvic tilt assessment, most of the participants, both boys and girls, showed posterior pelvic tilt for the L-V angle, regardless of BP. In contrast, for the L-H angle, girls showed a more neutral pelvis and boys more posterior pelvic tilt, but no association with BP were found. Similar results were found in the assessment of ROM, boys showed reduced mobility in hip extensor in a greater proportion than girls, but again no association were found with BP. However, for ROM of hip flexors, associations were found between a reduced ROM and BP. When analysing according to sex these associations only appeared in girls.

Concerning LBP, Table 18 shows the comparative analysis of the different continuous variables for the total sample of participants with and without LBP and according to sex. As with general BP, variables related to body weight showed significant differences between participants who had LBP and those who did not. In the spine assessment, no significant differences were found for any of the variables assessed. About ROM, differences were found again for PHE in total sample and girls. Finally, in the trunk muscle endurance tests, significant differences were found for SB test both in the total sample and among boys and girls with and without LBP, showing a worse performance those with LBP.

The distribution of the participants with and without LBP according to the classification of spinal alignment by evaluation position was similar to that found for BP in general, but in this case, no type of association was found to highlight. Concerning ROM classification, only association were found for hip flexors ROM and LBP. Participants with reduced hip flexors were associated with LBP ($\chi^2_{(1)}$=6.90, *p*=0.01), as well as girls with reduced ROM and LBP ($\chi^2_{(1)}$=4.54, *p*=0.03).

Table 17. Percentages (counts) of participants for each spinal curve by assessment position, pelvic tilt classification and prevalence of back pain by total sample and by sex

Variable	Classification	Total sample (n=252) No BP (n=181)	Total sample BP (n=71)	Total sample Chi-square	Male (n=122) No BP (n=87)	Male BP (n=35)	Male Chi-square	Female (n=130) No BP (n=94)	Female BP (n=36)	Female Chi-square
Thoracic curvature										
Standing position	Hypokyphosis	0.6% (1)	1.4% (1)	$\chi^2_{(2)}$=1.32	0% (0)	0% (0)	$\chi^2_{(2)}$=0.01	1.1% (1)	2.8% (1)	$\chi^2_{(2)}$=1.93
	Normal	58.6% (106)	53.5% (38)	*p*=0.47	50.6% (44)	51.4% (18)	*p*=0.93	66% (62)	55.6% (20)	*p*=0.36
	Hyperkyphosis	40.9% (74)	45.1% (32)		49.4% (43)	48.6% (17)		33% (31)	41.7% (15)	
Slump Sitting position	Hypokyphosis	0.6% (1)	1.4% (1)	$\chi^2_{(2)}$=1.23	0% (0)	0% (0)	$\chi^2_{(2)}$=0.01	1.1% (1)	2.8% (1)	$\chi^2_{(2)}$=1.59
	Normal	71.8% (130)	67.6% (48)	*p*=0.52	63.2% (55)	62.9% (22)	*p*=0.93	79.8% (75)	72.2% (26)	*p*=0.43
	Hyperkyphosis	27.6% (50)	31% (22)		36.8% (32)	37.% (13)		19.1% (18)	25% (9)	
Bending forward position	Hypokyphosis	0% (0)	0% (0)	$\chi^2_{(2)}$=0.05	0% (0)	0% (0)	$\chi^2_{(2)}$=0.10	0% (0)	0% (0)	$\chi^2_{(2)}$=0.32
	Normal	28.2% (51)	26.8% (19)	*p*=0.82	23% (20)	25.7% (9)	*p*=0.74	33% (31)	27.8% (10)	*p*=0.56
	Hyperkyphosis	71.8% (130)	73.2% (52)		77% (67)	74.3% (26)		67% (63)	72.2% (26)	
Lumbar curvature										
Standing position	Hypolordosis	5.5% (10)	9.9% (7)	$\chi^2_{(2)}$=1.67	9.2% (8)	11.4% (4)	$\chi^2_{(2)}$=0.27	2.1% (2)	8.3% (3)	$\chi^2_{(2)}$=2.71
	Normal	74.6% (135)	69% (49)	*p*=0.43	75.9% (66)	71.4% (25)	*p*=0.87	73.4% (69)	66.7% (24)	*p*=0.25
	Hyperlordosis	19.9% (36)	21.1% (15)		14.9% (13)	17.1% (6)		24.5% (23)	25% (9)	
Slump Sitting position	Lordosis	0% (0)	0% (0)	$\chi^2_{(2)}$=5.37	0% (0)	0% (0)	$\chi^2_{(2)}$=1.31	0% (0)	0% (0)	$\chi^2_{(2)}$=4.30
	Normal	60.6% (109)	76.1% (54)	*p*=0.03	52.3% (45)	65.7% (23)	*p*=0.25	68.1% (64)	86.1% (31)	*p*=0.04
	Hyperkyphosis	39.4% (71)	23.9% (17)		47.7% (41)	34.3% (12)		31.9% (30)	13.9% (5)	
Bending forward position	Hypokyphosis	3.9% (7)	0% (0)	$\chi^2_{(2)}$=8.77	4.6% (4)	0% (0)	$\chi^2_{(2)}$=4.10	3.2% (3)	0% (0)	$\chi^2_{(2)}$=3.71
	Normal	66.9% (121)	84.5% (60)	*p*=0.01	60.9% (53)	80% (28)	*p*=0.11	72.3% (68)	88.9% (32)	*p*=0.11
	Hyperkyphosis	29.3% (53)	15.5% (11)		34.5% (30)	20% (7)		24.5% (23)	11.1% (4)	
Pelvic tilt										
L-H angle	Neutral	43.6% (79)	45.1% (32)	$\chi^2_{(1)}$=0.04	33.3% (29)	25.7% (9)	$\chi^2_{(1)}$=0.67	53.2% (50)	63.9% (23)	$\chi^2_{(1)}$=1.21
	Posterior PT	56.4% (102)	54.9% (39)	*p*=0.83	66.7% (58)	74.3% (26)	*p*=0.41	46.8% (44)	36.1% (13)	*p*=0.27
L-V angle	Neutral	21% (38)	16.9% (12)	$\chi^2_{(1)}$=0.53	8% (7)	5.7% (2)	$\chi^2_{(1)}$=0.19	33% (31)	27.8% (10)	$\chi^2_{(1)}$=0.32
	Posterior PT	79% (143)	83.1% (59)	*p*=0.46	92% (80)	94.3% (33)	*p*=0.65	67% (63)	72.2% (26)	*p*=0.56
Range of Motion										
PHE_R	Normal	63.5% (115)	49.3% (35)	$\chi^2_{(1)}$=4.29	57.5% (50)	45.7% (16)	$\chi^2_{(1)}$=1.39	69.1% (65)	52.8% (19)	$\chi^2_{(1)}$=3.05
	Reduced	36.5% (66)	50.7% (36)	*p*=0.04	42.5% (37)	54.3% (19)	*p*=0.23	30.9% (29)	47.2% (17)	*p*=0.08
PHE_L	Normal	62.4% (113)	52.1% (37)	$\chi^2_{(1)}$=2.25	55.2% (48)	48.6% (17)	$\chi^2_{(1)}$=0.43	69.1% (65)	55.6% (20)	$\chi^2_{(1)}$=2.12
	Reduced	37.6% (68)	47.9% (34)	*p*=0.13	44.8% (39)	51.4% (18)	*p*=0.51	30.9% (29)	44.4% (16)	*p*=0.14
PHF_KE_R	Normal	41.4% (75)	33.8% (24)	$\chi^2_{(1)}$=1.24	28.7% (25)	20% (7)	$\chi^2_{(1)}$=0.98	53.2% (50)	47.2% (17)	$\chi^2_{(1)}$=0.37
	Reduced	58.6% (106)	66.2% (47)	*p*=0.26	71.3% (62)	80% (28)	*p*=0.32	46.8% (44)	52.8% (19)	*p*=0.54
PHF_KE_L	Normal	45.9% (83)	40.8% (29)	$\chi^2_{(1)}$=0.51	34.5% (30)	31.4% (11)	$\chi^2_{(1)}$=0.10	56.4% (53)	50% (18)	$\chi^2_{(1)}$=0.43
	Reduced	54.1% (98)	59.2% (42)	*p*=0.47	65.5% (57)	68.6% (24)	*p*=0.75	43.6% (41)	50% (18)	*p*=0.51

Note: L-H: lumbo-horizontal, L-V: lumbo-vertical, PT: pelvic tilt, PHF_KE (R-L): passive hip flexion with knee extension (right-left), PHE (R-L): passive hip extension (right-left)

Table 18. Descriptive values of anthropometric characteristics and "Fitness Postural" assessment (mean ± SD) for participants with or without low back pain by total sample and sex.

Variables	Total sample (n=252) No LBP (n=205)	Total sample (n=252) LBP (n=47)	Male (n=122) No LBP (n=99)	Male (n=122) LBP (n=23)	Female (n=130) No BP (n=106)	Female (n=130) LBP (n=24)
Anthropometric Characteristics						
Age (y)	10.9±1.2	11.1±1.2	11.1±1.4	11.2±1.5	10.7±0.9	10.9±0.9
Maturity offset	-1.6±1.2	-1.4±1.2	-2.3±1.1	-1.9±1.2	-1±0.9	-0.8±1
Body mass (kg)	42.9±10.9	47.5±12.2*	43.6±11	48.8±11.4*	42.4±10.7	46.4±13
Body height (cm)	145.6±9.1	147.3±10.7	146.6±9.8	148.7±10.2	144.7±8.5	145.9±11.1
BMI (kg/m^2)	20.1±3.9	21.7±3.9*	20.1±3.8	21.8±4*	20.1±4.1	21.5±3.9
Percentile	67±30.2	78.3±24.4*	67.3±30.3	79.8±23.8	66.8±30.1	76.9±25.3
Postural Fitness						
Spine Assessment						
SP thoracic (º)	38.9±9.1	38.5±10.8	40.8±9.1	40.9±9.7	37.1±8.8	36.2±11.6
SP lumbar (º)	33.6±9.5	34.3±10.2	31.2±9.3	33.4±10.5	35.8±9.2	35.2±9.9
SSP thoracic (º)	36.2±10.5	35.9±12.2	38.6±10.9	40±10.4	33.9±9.5	32±12.8
SSP lumbar (º)	11.5±11	9.4±9.2	13.4±12	11.3±9	9.7±9.8	7.6±9.1
L-H angle (º)	102.4±8.7	101.9±7.8	104.8±8.9	105±6.3	100±7.9	99±8.1
BFP thoracic (º)	48.8±11.9	47.9±10.2	50.5±10.8	48.8±11.3	47.3±12.6	47.1±9.1
BFP lumbar (º)	26.4±9.4	26±9.3	27..±9.4	27.9±6.7	24.8±9.1	24.4±6.7
L-V angle (º)	114.4±14.4	114.7±12.8	120.7±12.9	119.9±13	108.5±13.2	109.7±10.5
Range of Motion						
PHE-R (º)	17.9±8.4	14.3±7.9*	16.6±8.5	13.7±7.8	19±8.1	14.8±7.9*
PHE-L (º)	18±8.3	14.2±7.6*	16.9±8.3	13.9±7.1	19±8.2	14.4±8.2*
PHFKE-R (º)	74.2±13.4	72.4±10.8	69.3±9.8	69±9.1	78.7±14.6	75.6±11.4
PHFKE-L (º)	73.4±13.2	71.8±10.5	68.6±10.5	69±9.2	77.9±13.9	74.5±11.2
TT distance (cm)	-7.5±8.9	-7.8±8.2	-10.3±7.8	-9.9±7.6	-4.8±9	-5.9±8.5
Trunk Muscle Endurance						
Ito test (s)	86.4±75.3	72.3±54.9	87.1±85.8	82.8±64.6	85.7±64.8	63.1±44.1*
BS test (s)	131.7±72.7	119.1±63.3	124.8±76.5	104.1±54.2	138.2±68.7	133.5±69
SB-R test (s)	46±25.8	34.8±21.7*	50±27.6	37.3±25*	42.5±23.6	32.5±18.4*
SB-L test (s)	47.2±27.3	35.7±20.8*	53±29.9	37±25.4*	41.9±23.7	34.4±15.9*

Note: y: years, kg: kilograms, cm: centimetres, BMI: body mass index, º: degrees, SP: standing position, SSP: slump sitting position, L-H: lumbo-horizontal, BFP: bending forward position, L-V: lumbo-vertical, PHE (R-L): passive hip extension (right-left), PHFKE (R-L): passive hip flexion with the knee extended (right-left), TT: "Toe Touch" test, BS: Biering-Sorensen, s: seconds, SB (R-L): Side Bridge (right-left). *Significant differences between those with and without LBP ($p<0.05$).

3.4. Discussion

Given that the prevalence of BP increases from childhood to adolescence, more specifically at around 11-12 years or even more after the onset of puberty [10,14,58], in the first part of this study the anthropometric characteristics and "Postural Fitness" (sagittal spinal alignment, mobility of the spine, range of motion (ROM) of the hip and trunk muscle endurance) during childhood and adolescence were analysed taking as reference the maturational stage of the participants. However, it was not possible to perform the analyses according to the state of maturation because the age range of the participants did not allow the inclusion of a homogeneous number of students in each maturational stage group (pre-, circa- and post-PHV) for both sexes (104 boys and 67 girls in pre-PHV, and 18 boys and 63 girls in circa-PHV). Indeed, no student was found in this sample that had already completed pubertal development (post-PHV).

The first hypothesis raised at the beginning of the study was confirmed because sex-related differences were found in all tests of the “Postural Fitness” protocol. This is the first study that performs a complete assessment of the sagittal morphotype of the spine, the hip ROM and trunk muscle endurance in children and adolescents, to our knowledge.

For the spine assessment, boys showed greater thoracic kyphosis compared to girls in the 3 measured positions, and girls presented a more lordotic lumbar curvature in standing and lower lumbar kyphosis and more neutral pelvic tilt in slump sitting and bending forward positions. Hence, when determining the “Sagittal integral Morphotype” for the thoracic curve, boys showed a tendency towards a more “hyperkyphotic” morphotype, while the “normal” morphotype predominated among girls. For the lumbar curve, girls showed a more “hyperlordotic” morphotype in standing but when establishing the “Sagittal Integral Morphotype”, the most common was the “normal” morphotype for both boys and girls. In the few studies that have assessed spinal alignment in the three defined positions, the results are similar [33,51,52]. In fact, in the only study [33] where the comprehensive diagnosis has been determined also found a higher percentage of children with a “functional thoracic hyperkyphosis” morphotype (36.8% in their study vs 40.1% in this study). Concerning the lumbar diagnosis, in the study of Santonja-Medina [33], almost all schoolchildren presented a “functional lumbar hyperkyphosis” morphotype (82.4%). Conversely, in the present work, the most common lumbar morphotype was “normal” (36.9%) followed by “functional hyperkyphosis” (33.7%). This difference may be since in the present study the “normal” lumbar morphotype was more prevalent among girls (40.8%) and the “functional hyperkyphosis” morphotype among boys (38.5%).

Sex-related differences during the study of spinal alignment have been seen in numerous studies, in which girls showed a “hyperlordotic” posture due to increased lumbar lordosis and boys a “sway” posture [59–63]. Also, both Schlosser et al. and Cil et al. found that dorsal and lumbar curvatures increase as schoolchildren grow and the differences found in curvatures between boys and girls were related to maturation status [61,63]. On the one hand, it was seen that the increase in curvatures was not linear, lumbar lordosis increased considerably until the age of 10-12 years, while dorsal kyphosis began to grow significantly from the age of 13. Thus, lumbar lordosis completes its growth several years before thoracic kyphosis [57]. On the other hand, the thoracic curvature was similar between boys and girls before and after the growth spurt, but during the onset of puberty, boys presented greater dorsal kyphosis. For lumbar lordosis, they did not find these differences in terms of pubertal development and sex, simply that the curvature was greater among girls [59]. It has been a shame not being able to analyse the "Sagittal Integral Morphotype" according to the maturational stage and check if it also changes according to the stage of the participants.

Sex-related differences were also found for hip ROM and trunk mobility, as in most of the studies that assess the ROM in children and adolescents, in the present study the girls showed a greater range in all the tests performed [64–68]. A possible explanation may be due to hormonal differences between boys and girls, since girls secrete more estrogen, and have higher levels of relaxin, and anatomical differences such as in the lumbo-pelvic region [46,69,70]. Furthermore, a reduced ROM in hip extensors (PHF_KE test) was found in most participants, it was more common among boys, as well as a posterior pelvic tilt. As a result of these findings, it is important to note that the muscles that originate or insert in the pelvis, depending on how they are found (normal or reduced ROM, strong or weak muscles), could generate changes in the pelvis tilt and, in turn, in the spine alignment.

The last part of the "Postural Fitness" protocol consists of the assessment of trunk muscles endurance. In this study, boys showed better performance in the test for the lateral-flexor muscles (SB test) and girls for the trunk extensors (BS test). Previous research also found sex-related differences for trunk endurance performance in children and adolescents [71–77]. To understand the higher performance of girls in the BS test, some authors usually refer to the fact that girls have greater lumbar lordosis than boys [52,78–80]. They suggest that this greater lumbar curvature would allow for higher mechanical advantage of the erector spinae muscles [81]. Other possible explanations could be the different geometry of female torso from male [82], as well as the presence of a greater number of type I fibres in the lumbar region [83].

As a result of the differences found between boys and girls for “Postural Fitness”, it is likely that there are also sex-related differences between the triggers of BP in children and adolescents. Therefore, in the second part of this study, the possible associations between these physical factors and BP (determine if these physical factors are possible triggers of BP) were explore.

The 1-year BP prevalence found in the present study (28.2%) was similar to that found in previous studies. For example, in Spanish elementary students (8-12 years) the mean prevalence was 22.3%, although in Spanish high school students (12-17 years) the prevalence was higher (55.1%) [50,52]. Even though no differences were found between boys and girls for the prevalence of BP (28.7% in boys and 27.7% in girls), previous studies have shown that girls have a higher risk of BP and that pain in girls is associated with chronological age and maturity status [13,14].

It should be noted that for the anthropometric variables, differences were only found between boys with BP and without BP. Among boys with BP or LBP, an increase of body mass and BMI were found compared with their peers asymptomatic. In a recent longitudinal study, the association of BP with weight status was studied during childhood (from 4 years to 15 years), finding that individuals classified as overweight or obese had a higher risk of developing LBP at any site. Furthermore, adjusting by age, sex, socioeconomic level and nationality, a child with obesity showed a 34% higher risk of developing BP before the age of 15 [27]. Some authors suggest that being overweight or obese during childhood could increase the mechanical load on the spine by increasing compressive forces on the immature structures of the spine and, in turn, interfere with the nutrition of the intervertebral disc and increase the risk of LBP [84–86].

For the spine assessment and BP, associations were found between BP and the participants who presented a "normal" morphotype in slump sitting and bending forward positions, as well as between participants without BP and those who presented a "hyperkyphosis" morphotype in those positions, this association appearing in girls when analysed by sex. These results coincide with other studies where those who did not suffer from BP had greater lumbar kyphosis or a greater posterior pelvic tilt in slump sitting position than those with BP [51]. On the other hand, O'Sullivan et al. found no association between BP made worse by sitting and the degree of slump sitting [87].

However, these results do not indicate that greater lumbar kyphosis in slump sitting or bending forward position can prevent BP. Those who presented a hyperkyphotic morphotype in sitting and bending forward had a mean lumbar kyphosis of 20.5º in sitting and 37.1º in bending forward, which could be categorized as slight lumbar hyperkyphosis (16º to 25º in sitting and 31º to 40º for bending forward) [51,88]. On the other hand, most bad postural habits are established during growth and pubertal development in children and adolescents, causing postural changes, particularly frontal and sagittal changes in the spine, which can also be related to sitting posture, either when attending school or to watching TV [89]. Puberty is a vulnerable period for musculoskeletal disorders due to the occurrence of a wide variation in growth and development. Therefore, early correction of bad postural attitudes is essential, because it enables a better prognosis and treatment perspectives [89].

Concerning ROM, girls with BP or LBP showed a lower ROM in hip flexor muscles and in boys, a tendency toward significance was observed. When classifying hip ROM as normal or reduced, an association was also found between having a reduced ROM in hip flexor muscles and BP or LBP in girls. The iliopsoas is the main hip flexor and a reduced ROM, among other factors, could be due to the sitting posture maintained by students during their school and after school time [46,69]. Also, its multi-articular condition, its diversity of functions and its tonic-postural character, make it a muscle with a tendency to shorten [65,66]. For example, students enter a primary educational centre and until they leave their studies, they spend a lot of time sitting in the school rigid chairs which will lead to different muscular adaptations. The hip flexor muscles will be in a shortened posture every time we are in hip flexion. Schoolchildren adopt this posture for most of the time, which causes them to slowly and persistently shorten the iliopsoas [65,66]. If this shortening is “permanent” due to its frequency and occurs “early”, starting in childhood, it will slowly and progressively modify the osteoarticular structures of the spine [64,88,90].

The scientific evidence in terms of hip ROM seems more complex as the measurement methods differ substantially across studies as do the ages of explored samples and their findings [19]. Some researchers have reported that tightness of the hip flexor muscles [91], and decreased ROM in hamstring were associated with LBP in adolescents [92–94]. On the other hand, two studies by Sjölie [95] and Harreby et al. [96] that measured hamstrings ROM via knee extension deficit in 90 degrees of hip flexion (popliteal angle) found contradicting results concerning ROM and BP. In the present study, for example, no association were found between the ROM of hip extensor muscles and BP and LBP, neither for boys nor for girls. Gonzalez-Galvez et al. [97] did not find differences in the hamstrings ROM between the group with BP and the group without BP in adolescents and using the same test as in the present study (PHF_KE).

Finally, the assessment of trunk muscle endurance revealed associations between BP and LBP and the Ito test for girls, and between LBP and the SB test for girls and boys. Participants who reported having BP or LBP performed poorly compared to their asymptomatic peers. There are different research studies with children and adolescents where associations between trunk flexor endurance and BP were also found [92,94,98]. However, direct comparisons cannot be made because the field-based tests used in previous studies are different from those of the present study. On the other hand, it should be noted that that other studies found no association between trunk flexors endurance and BP [76,99]. In relation to trunk extensor endurance, there is a greater consensus regarding the field-based test used (BS test), and most studies find associations between BP and poor performance in the test. In the present work, no associations were found, although the test duration was lower in the participants with BP and LBP. In line with the present study, Smith et al. found no association between LBP and trunk extensor endurance in 14-year-old adolescents [100], in the follow-up study for those same students they found an association with the LBP at age 17 and trunk extensor endurance [99]. Therefore, the participants of the present study could show this association as pubertal development occurs since most of them were in the prepubertal phase. Furthermore, it has been seen that the prevalence of BP increases after 12 years (coinciding with the growth spurt in girls) [4,13,14,101], and that the differences in trunk muscle endurance are accentuated from the period of pubertal onset and when maturation is complete [72,76,77]. Concerning trunk lateral-flexor muscles endurance, no studies have been found that assess this musculature in the general child-adolescent population. However, in several studies with children and adolescent athletes, associations similar to those of the present work have been found, where athletes with BP showed an association with lower performance in the SB test [102–104].

The results of the present study concerning BP confirm some of the hypotheses raised at the beginning of the research. Participants with BP had reduced ROM in the hip flexors and worse endurance of the trunk lateral-flexor muscles, but they did not show worse spinal alignment. The physical triggers for BP were different according to sex, finding associations between boys with BP and a higher BMI and between girls with BP and a reduced ROM in the hip flexors and a worse resistance in the lateral flexors of the trunk.

Some limitations of the study should be pointed out. First, it is a cross-sectional study that only shows the association between possible risk factors for BP, so it is not possible to demonstrate a cause-effect relationship. Another limitation is the recall period analysed, the use of 1-year prevalence could cause a great loss of information because it depends on the ability of the participants to recall pain. Therefore, for future research, the lifetime prevalence will be used together with the 1-month prevalence or 1-week. Furthermore, the differences in the number of participants in the maturational stage groups and the absence of post-PHV participants prevented the analysis of "Postural Fitness" and BP based on the maturational stage. Lastly, the fact that the sample belongs to the same Spanish region does not allow the results to be extrapolated to other areas of Spain or other countries, since they are influenced by environmental, social, cultural and genetic factors specific to each site. However, the results of the study could be used within the community to promote prevention programs among educational centres.

To know how the different physical factors progress throughout childhood and adolescence, as well as the prevalence of BP, future research studies should include participants from 9 years to 18 years, to carry out the analyses according to the stage maturation and sex, since boys complete their maturation process later than girls (about 2 years apart). In this way, more effective preventive programs could be carried out. In addition, periodic screening of the spine is recommended in children from the beginning of the compulsory school stage, since children of this age represent a critical group for the development of spinal deformities and postural disorders [51,105].

3.5. Conclusion

In summary, the present study showed that children and adolescents in pre- and circa-PHV presented sex-related differences in the assessment of "Postural Fitness" protocol. Boys presented a morphotype more "hyperkyphotic" along with a greater posterior pelvic tilt and girls a morphotype more "hyperlordotic" and more neutral pelvic tilt. On the other hand, boys had a hip ROM reduced and more endurance in the lateral-flexor muscles. Girls, however, showed more hip ROM and more extensor muscle endurance.

Concerning BP, association were found between BP in boys and higher IMC, and between BP in girls and reduced ROM in hip flexor muscles and poor trunk flexor endurance. When analysing the associations between "Postural Fitness" and LBP, boys with LBP showed an association with a higher BMI and a worse trunk lateral-flexors endurance, and girls with LBP showed association with a reduced ROM in hip flexor muscles and a worse endurance for trunk flexor and lateral-flexor muscles.

Hip ROM and trunk muscle endurance are modifiable risk factors whose optimization can generate benefits in the musculoskeletal health of children and adolescents, taking into account always their sex and biological maturity, as well as the alignment of the spine, at the time to carry out educational prevention programs.

3.6. References

1. Hill, J.J.; Keating, J.L. Risk factors for the first episode of low back pain in children are infrequently validated across samples and conditions: a systematic

review. *J. Physiother.* 2010, *56*(4), 237–44, doi: 10.1016/S1836-9553(10)70006-0.

2. Kamper, S.; Yamato, T.P.; Williams, C.M. The prevalence, risk factors, prognosis and treatment for back pain in children and adolescents: An overview of systematic reviews. *Best Pract. Res. Clin. Rheumatol.* 2016, *30*(6), 1021–36, doi: 10.1016/j.berh.2017.04.003.
3. Junge, T.; Wedderkopp, N.; Boyle, E.; Kjaer, P. The natural course of low back pain from childhood to young adulthood - a systematic review. *Chiropr. Man. Ther.* 2019, *27*(1), 1–10, doi: 10.1186/s12998-018-0231-x.
4. Calvo-Muñoz, I.; Gómez-Conesa, A.; Sánchez-Meca, J. Prevalence of low back pain in children and adolescents: A meta-analysis. *BMC Pediatr.* 2013, *13*(14), 1–12, doi: 10.1186/1471-2431-13-14.
5. Jeffries, L.; Milanese, S.F.; Grimmer-Somers, K.A. Epidemiology of Adolescent Spinal Pain: A Systematic Overview of the Research Literature. *Spine (Phila Pa 1976).* 2007, *32*(23), 2630–7, doi: 10.1097/BRS.0b013e318158d70b.
6. Wirth, B.; Potthoff, T.; Rosser, S.; Humphreys, B.K.; De Bruin, E.D. Physical risk factors for adolescent neck and mid back pain: A systematic review. *Chiropr Man Ther.* 2018, *26*(1), 1–10, doi: 10.1186/s12998-018-0206-y.
7. Dionne, C.E.; Dunn, K.M.; Croft, P.R.; Nachemson, A.L.; Buchbinder, R.; Walker, B.F.; et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine (Phila Pa 1976).* 2008, *33*(1), 95–103, doi: 10.1097/BRS.0b013e31815e7f94.
8. Briggs, A.M.; Smith, A.J.; Straker, L.M.; Bragge, P. Thoracic spine pain in the general population: Prevalence, incidence and associated factors in children, adolescents and adults. A systematic review. *BMC Musculoskelet. Disord.* 2009, *10*(1), 77, doi: 10.1186/1471-2474-10-77.
9. Scarabottolo, C.; Pinto, R.; Oliveira, C.; Zanuto, E.; Cardoso, J.; Christofaro, D. Back and neck pain prevalence and their association with physical inactivity domains in adolescents. *Eur. Spine J.* 2017, *26*(9), 2274–80, doi: 10.1007/s00586-017-5144-1.
10. Kamper, S.; Henschke, N.; Hestbaek, L.; Dunn, K.; Williams, C. Musculoskeletal pain in children and adolescents. *Brazilian J. Phys. Ther.* 2016, *20*(3), 275–84, doi: 10.1590/bjpt-rbf.2014.0149.
11. Kjaer, P.; Wedderkopp, N.; Korsholm, L.; Leboeuf-Yde, C. Prevalence and

tracking of back pain from childhood to adolescence. *BMC Musculoskelet. Disord.* 2011, *12*(98), 1–11, doi: 10.1186/1471-2474-12-98.

12. Wedderkopp, N.; Andersen, L.B.; Froberg, K.; Leboeuf-Yde, C. Back pain reporting pattern in a Danish population-based sample of children and adolescents. *BMC Musculoskelet. Disord.* 2001, *6*(52), 1–5, doi: 10.1186/1471-2474-6-52.
13. Beynon, A.M.; Hebert, J.J.; Lebouef-Yde, C.; Walker, B.F. Potential risk factors and triggers for back pain in children and young adults. A scoping review, part I: Incident and episodic back pain. *Chiropr. Man. Ther.* 2019, *27*(58), 1–7, doi: 10.1186/s12998-019-0280-9.
14. Lardon, A.; Leboeuf-Yde, C.; Le Scanff, C.; Wedderkopp, N. Is puberty a risk factor for back pain in the young? A systematic critical literature review. *Chiropr. Man. Ther.* 2014, *22*(1), 1–12, doi: 10.1186/s12998-014-0027-6.
15. Dunn, K.M.; Hestbaek, L.; Cassidy, J.D. Low back pain across the life course. *Best Pract. Res. Clin. Rheumatol.* 2013, *27*(5), 591–600, doi: 10.1016/j.berh.2013.09.007.
16. Hestbaek, L.; Leboeuf-Yde, C.; Kyvik, K.O.; Manniche, C. The course of low back pain from adolescence to adulthood: Eight-year follow-up of 9600 twins. *Spine (Phila Pa 1976).* 2006, *31*(4), 468–72, doi: 10.1097/01.brs.0000199958.04073.d9.
17. Jones, G.T.; Watson, K.D.; Silman, A.J.; Symmons, D.P.M.; Macfarlane, G.J. Predictors of low back pain in British schoolchildren: A population-based prospective cohort study. *Pediatrics.* 2003, *111*(4), 822–8, doi: 10.1542/peds.111.4.822.
18. Kovacs, F.; Gestoso, M.; Gil del Real, M.; Lopez, J.; Mufraggi, N.; Mendez, J. Risk factors for non-specific low back pain in schoolchildren and their parents: A population based study. *Pain.* 2003, *103*(3), 259–68, doi: 10.1016/S0304-3959(02)00454-2.
19. Potthoff, T.; De Bruin, E.D.; Rosser, S.; Humphreys, B.K.; Wirth, B. A systematic review on quantifiable physical risk factors for non-specific adolescent low back pain. *J. Pediatr. Rehabil. Med.* 2018, *11*(2), 79–94, doi: 10.3233/PRM-170526.
20. Calvo-Muñoz, I.; Kovacs, F.M.; Roqué, M.; Gago, I.; Seco, J. Risk Factors for Low Back Pain in Childhood and Adolescence: A Systematic Review. *Clin. J*

Pain. 2018, 34(*5*), 468–484, doi: 10.1097/AJP.0000000000000558.

21. Beynon, A.M.; Hebert, J.J.; Lebouef-Yde, C.; Walker, B.F. Potential risk factors and triggers for back pain in children and young adults. A scoping review, part II: unclear or mixed types of back pain. *Chiropr. Man. Ther.* 2019, *27*(61), 1–12, doi: 10.1186/s12998-019-0281-8.
22. Beales, D.J.; Smith, A.J.; O'Sullivan, P.B.; Straker, L.M. Low Back Pain and Comorbidity Clusters at 17 Years of Age: A Cross-sectional Examination of Health-Related Quality of Life and Specific Low Back Pain Impacts. *J. Adolesc. Health*. 2012, *50*(5), 509–16, doi: 10.1016/j.jadohealth.2011.09.017.
23. Watson, K.D.; Papageorgiou, A.C.; Jones, G.T.; Taylor, S.; Symmons, D.P.M.; Silman, AJ.; et al. Low back pain in schoolchildren: the role of mechanical and psychosocial factors. *Arch. Dis. Child.* 2003, *88*(1), 12–7, doi: 10.1136/adc.88.1.12.
24. Mikkonen, P.; Heikkala, E.; Paananen, M.; Remes, J.; Taimela, S.; Auvinen, J.; et al. Accumulation of psychosocial and lifestyle factors and risk of low back pain in adolescence: a cohort study. *Eur. Spine J.* 2016, *25*(2), 635–42, doi: 10.1007/s00586-015-4065-0.
25. Alzahrani, H.; Mackey, M.; Stamatakis, E.; Zadro, J.R.; Shirley, D. The association between physical activity and low back pain: a systematic review and meta-analysis of observational studies. *Sci. Rep.* 2019, *9*(1), 1–10, doi: 10.1038/s41598-019-44664-8.
26. Calvo-Muñoz, I.; Kovacs, F.M.; Roqué, M.; Seco-Calvo, J. The association between the weight of schoolbags and low back pain among schoolchildren: A systematic review, meta-analysis and individual patient data meta-analysis. *Eur. J. Pain (United Kingdom)*. 2020, *24*(1), 91–109, doi: 10.1002/ejp.1471.
27. Palmer, A.; Poveda, J.; Martinez-Laguna, D.; Reyes, C.; de Bont, J.; Silman, A.; et al. Childhood overweight and obesity and back pain risk: a cohort study of 466997 children. *BMJ Open*. 2020, *10*(e036023), doi: 10.1136/bmjopen-2019-036023.
28. Dolphens, M.; Vansteelandt, S.; Cagnie, B.; Vleeming, A.; Nijs, J.; Vanderstraeten, G.; et al. Multivariable modeling of factors associated with spinal pain in young adolescence. *Eur. Spine J.* 2016, *25*(9), 2809–21, doi: 10.1007/s00586-016-4629-7.
29. Fogel, B.; Greenberg, D. *Psychiatric Care of the Medical Patient Psychiatric*. In:

Care of the Medical Patient. Oxford, UK: Oxford Uni, 2015.,

30. Mirwald, R.L.; Baxter-Jones, A.D.G.; Bailey, D.A.; Beunen, G.P. An assessment of maturity from anthropometric measurements. *Med. Sci. Sports Exerc.* 2002, *34*(4), 689–94, doi: 10.1097/00005768-200204000-00020.
31. Radnor, J.M.; Oliver, J.L.; Waugh, C.M.; Myer, G.D.; Lloyd, R.S. The influence of maturity status on muscle architecture in school-aged boys. *Pediatr. Exerc. Sci.* 2020, *32*(2), 89–96, doi: 10.1123/pes.2019-0201.
32. Robles-Palazón, F.J.; Ayala, F.; Cejudo, A.; De Ste Croix, M.; Sainz de Baranda, P.; Santonja, F. Effects of Age and Maturation on Lower Extremity Range of Motion in Male Youth Soccer Players. *J. Strength Cond. Res.* 2020, *Publish Ah*, doi: 10.1519/jsc.0000000000003642.
33. Santonja-Medina, F.; Collazo-Diéguez, M.; Martínez-Romero, M.T.; Rodríguez-Ferrán, O.; Aparicio-Sarmiento, A.; Cejudo, A.; et al. Classification system of the sagittal integral morphotype in children from the isquios programme (Spain). *Int. J. Environ. Res. Public Health.* 2020, *17*(7), 1–15, doi: 10.3390/ijerph17072467.
34. Cejudo, A.; Sainz de Baranda, P.; Ayala, F.; De Ste Croix, M.; Santonja-Medina, F. Assessment of the Range of Movement of the Lower Limb in Sport: Advantages of the ROM-SPORT I Battery. *Int. J. Environ. Res. Public Health.* 2020, *17*(20), 7606, doi: 10.3390/ijerph17207606.
35. Ito, T.; Shirado, O.; Suzuki, H.; Takahashi, M.; Kaneda, K.; Strax, T.E. Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. *Arch. Phys. Med. Rehabil.* 1996, *77*, 75–9, doi: 10.1016/S0003-9993(96)90224-5.
36. McGill, S.M.; Childs, A.; Liebenson, C. Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. *Arch. Phys. Med. Rehabil.* 1999, *80*(8), 941–4, doi: 10.1016/S0003-9993(99)90087-4.
37. Biering-Sorensen, F. Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine (Phila Pa 1976).* 1984, *9*(2), 106–19.
38. Ginés-Díaz, A.; Martínez-Romero, M.T.; Cejudo, A.; Aparicio-Sarmiento, A.; Sainz de Baranda, P. Sagittal Spinal Morphotype Assessment in Dressage and Show Jumping Riders. *J. Sport Rehabil.* 2019, *29*(5), 533–40, doi: 10.1123/jsr.2018-0247.
39. Sainz de Baranda, P.; Rodríguez-García, P.L.; Santonja, F. Efectos sobre la

disposición sagital del raquis de un programa de Educación Postural en Educación Física de Primaria. *Apunt Educ. Física y Deport.* 2010, *4*(102), 16–21.

40. Sainz de Baranda, P.; Santonja, F. Valoración de la disposición sagital del raquis en gimnastas especialistas en trampolín. *Int. J. Sport Sci.* 2009, *5*(16), 21–33, doi: 10.5232/ricyde2009.016.02.
41. Sanz-Mengibar, J.M.; Sainz de Baranda, P.; Santonja, F. Training intensity and sagittal curvature of the spine in male and female artistic gymnasts. *J. Sports Med. Phys. Fitness.* 2018, *58*(4), 465–71, doi: 10.23736/S0022-4707.17.06880-3.
42. Sainz De Baranda, P.; Rodriguez-Iniesta, M.; Ayala, F.; Santonja, F.; Cejudo, A. Determination of the criterion-related validity of hip joint angle test for estimating hamstring flexibility using a contemporary statistical approach. *Clin J Sport Med.* 2014, *24*(4), 320–5, doi: 10.1097/JSM.0000000000000079.
43. Ayala, F.; Sainz De Baranda, P.; Cejudo, A.; Santonja, F. Pruebas angulares de estimación de la flexibilidad isquiosural: descripción de los procedimientos exploratorios y valores de referencia. *Rev. Andaluza Med. del Deport.* 2013, *6*(3), 120–8, doi: 10.1016/S1888-7546(13)70046-7.
44. Santonja, F.; Andújar, P.; Martínez, I. Ángulo lumbo-horizontal y valoración de repercusiones del Síndrome de Isquiosurales Cortos. *Apunt Educ. Fis. y Deport.* 1994, (31), 103–11.
45. Ayala, F.; Sainz de Baranda, P.; De Ste Croix, M.; Santonja, F. Reproducibility and criterion-related validity of the sit and reach test and toe touch test for estimating hamstring flexibility in recreationally active young adults. *Phys. Ther. Sport.* 2012, *13*(4), 219–26, doi: 10.1016/j.ptsp.2011.11.001.
46. Perret, C.; Poiraudeau, S.; Fermanian, J.; Lefèvre Colau, M.M.; Mayoux Benhamou, M.A.; Revel, M. Validity, reliability, and responsiveness of the fingertip-to-floor test. *Arch. Phys. Med. Rehabil.* 2001, *82*(11), 1566–70, doi: 10.1053/apmr.2001.26064.
47. Juan-Recio, C.; Barbado, D.; López-Valenciano, A.; Vera-García, F.J. Test de campo para valorar la resistencia de los músculos del tronco. *Apunt Educ. Fis. y Deport.* 2014, *3*(117), 59–68, doi: 10.5672/apunts.2014-0983.es.(2014/3).117.06.
48. Aparicio-Sarmiento, A.; Rodríguez-Ferrán, O.; Martínez-Romero, M.T.; Cejudo, A.; Santonja, F.; Sainz de Baranda, P. Back pain and knowledge of back care related to physical activity in 12 to 17 year old adolescents from the region of Murcia (Spain): ISQUIOS programme. *Sustain.* 2019, *11*(19), 5249, doi:

10.3390/su11195249.

49. Sainz de Baranda, P.; Cejudo, A.; Martínez-Romero, M.T.; Aparicio-Sarmiento, A.; Rodríguez-Ferrán, O.; Collazo-Diéguez, M.; et al. Sitting posture, sagittal spinal curvatures and back pain in 8 to 12-year-old children from the region of murcia (Spain): ISQUIOS programme. *Int. J. Environ. Res. Public Health*. 2020, *17*(7), 1–18, doi: 10.3390/ijerph17072578.

50. Sainz de Baranda, P.; Andújar, P.; Collazo-Diéguez, M.; Pastor, A.; Santonja-Renedo, F.; Martínez-Romero, MT.; et al. Sagittal standing spinal alignment and back pain in 8 to 12-year-old children from the Region of Murcia, Spain: ISQUIOS programme. *J. Back Musculoskelet. Rehabil.* 2020, *33*(6), 1003–14, doi: 10.3233/BMR-191727.

51. Martínez-Crespo, G.; Rodríguez-Piñero, M.; López-Salguero, A.I.; Zarco-Periñan, M.J.; Ibáñez-Campos, T.; Echevarría-Ruiz de Vargas, C.; et al. Dolor de espalda en adolescentes: prevalencia y factores asociados. *Rehabilitación*. 2009, *43*(2), 72–80, doi: 10.1016/S0048-7120(09)70773-X.

52. Rodríguez-Oviedo, P.; Ruano-Ravina, A.; Pérez-Ríos, M.; García, F.B.; Gómez-Fernández, D.; Fernández-Alonso, A.; et al. School children’s backpacks, back pain and back pathologies. *Arch. Dis. Child*. 2012, *97*(8), 730–2, doi: 10.1136/archdischild-2011-301253.

53. Hopkins, W.G.; Marshall, S.W.; Batterham, A.M.; Hanin, J. Progressive Statistics for Studies in Sports Medicine and Exercise Science. *Med. Sci. Sports Exerc*. 2009, *41*(1), 3–12, doi: 10.1249/MSS.0b013e31818cb278.

54. Janssens, K.A.M.; Rosmalen, J.G.M.; Ormel, J.; Verhulst, F.C.; Hunfeld, J.A.M.; Mancl, L.A.; et al. Pubertal status predicts back pain, overtiredness, and dizziness in american and dutch adolescents. *Pediatrics*. 2011, *128*(3), 553–9, doi: 10.1542/peds.2010-2364.

55. Araújo, F.A.; Martins, A.; Alegrete, N.; Howe, L.D.; Lucas, R. A shared biomechanical environment for bone and posture development in children. *Spine J*. 2017, *17*(10), 1426–34, doi: 10.1016/j.spinee.2017.04.024.

56. Dolphens, M.; Cagnie, B.; Vleeming, A.; Vanderstraeten, G.; Danneels, L. Gender differences in sagittal standing alignment before pubertal peak growth: The importance of subclassification and implications for spinopelvic loading. *J. Anat*. 2013, *223*(6), 629–40, doi: 10.1111/joa.12119.

57. Schlosser, T.P.C.; Vincken, K.L.; Rogers, K.; Castelein, R.M.; Shah, S.A.

Natural sagittal spino-pelvic alignment in boys and girls before , at and after the adolescent growth spurt. *Eur. Spine j.* 2015, *24*(6), 1158–67, doi: 10.1007/s00586-014-3536-z.

58. Smith, A.; O'Sullivan, P.; Straker, L. Classification of sagittal thoraco-lumbo-pelvic alignment of the adolescent spine in standing and its relationship to low back pain. *Spine (Phila Pa 1976).* 2008, *33*(19), 2101–7, doi: 10.1097/BRS.0b013e31817ec3b0.
59. Cil, A.; Yazici, M.; Uzumcugil, A.; Kandemir, U.; Alanay, A.; Alanay, Y.; et al. The evolution of sagittal segmental alignment of the spine during childhood. *Spine (Phila Pa 1976).* 2004, *30*(1), 93–100, doi: 10.1097/01.brs.0000149074.21550.32.
60. Collazo-Diéguez, M. *Morfotipos Sagitales del Raquis en Población Escolar de la Región de Murcia* [Thesis]. Universidad de Murcia; 2015.
61. Wirth, B.; Knecht, C.; Humphreys, K. Spine day 2012: Spinal pain in Swiss school children- epidemiology and risk factors. *BMC Pediatr.* 2013, *13*(159), 1–10, doi: 10.1186/1471-2431-13-159.
62. Alricsson, M.; Bjorklund, G.; Cronholm, M.; Olsson, O.; Viklund, P.; Svantesson, U.; et al. Spinal alignment, mobility of the hip and thoracic spine and prevalence of low back pain in young elite cross-country skiers. *J. Exerc. Rehabil.* 2016, *12*(1), 21–8, doi: 10.12965/jer.150255.
63. Aartun, E.; Hartvigsen, J.; Hestbaek, L. Validity of Commonly Used Clinical Tests to Diagnose and Screen for Spinal Pain in Adolescents: A School-Based Cohort Study in 1300 Danes Aged 11-15 Years. *J. Manipulative Physiol. Ther.* 2016, *39*(2), 76–87, doi: 10.1016/j.jmpt.2016.01.007.
64. Sankar, W.N.; Laird, C.T.; Baldwin, K.D. Hip range of motion in children: What is the norm? *J. Pediatr. Orthop.* 2012, *32*(4), 399–405, doi: 10.1097/BPO.0b013e3182519683.
65. Martinez-Garcia, A.C. *Efectos de un Programa de Educación Postural sobre el Morfotipo Sagital del Raquis, la Extensibilidad de la Musculatura Isquiosural y Psoas Iliaco y la Resistencia Muscular Abdominal y Lumbar en Escolares de Educación Secundaria* [Thesis]. Universidad de Murcia; 2013.
66. Espada, D.R.; Montesinos, J.L.G.; Vicente, J.M. Diferencias en las amplitudes articulares entre varones y mujeres en edad escolar. *Apunt Med. l'Esport*. 2007, *42*(153), 13–25, doi: 10.1016/s1886-6581(07)70032-4.

67. Shultz, S.J.; Nguyen, A.D.; Schmitz, R.J. Differences in lower extremity anatomical and postural characteristics in males and females between maturation groups. *J. Orthop Sports Phys. Ther.* 2008, *38*(3), 137–49, doi: 10.2519/jospt.2008.2645.

68. Dejanovic, A.; Cambridge, E.D.J.; McGill, S. Isometric torso muscle endurance profiles in adolescents aged 15-18: normative values for age and gender differences. *Ann. Hum. Biol.* 2014, *41*(2), 153–8, doi: 10.3109/03014460.2013.837508.

69. Dejanovic, A.; Harvey, E.P.; McGill, S.M. Changes in Torso Muscle Endurance Profiles in Children Aged 7 to 14 Years: Reference Values. *Arch. Phys. Med. Rehabil.* 2012, *93*(12), 2295–301, doi: 10.1016/j.apmr.2011.12.023.

70. Moya-Ramón, M.; Juan-Recio, C.; Lopez-Plaza, D.; Vera-Garcia, F.J. Dynamic trunk muscle endurance profile in adolescents aged 14-18: Normative values for age and gender differences. *J. Back Musculoskelet. Rehabil.* 2018, *31*(1), 155–62, doi: 10.3233/BMR-169760.

71. Arampatzis, A.; Frank, J.; Laube, G.; Mersmann, F. Trunk muscle strength and lumbo-pelvic kinematics in adolescent athletes: Effects of age and sex. *Scand J. Med. Sci. Sport.* 2019, *29*(11), 1691–8, doi: 10.1111/sms.13503.

72. Mueller, J.; Mueller, S.; Stoll, J.; Baur, H.; Mayer, F. Trunk extensor and flexor strength capacity in healthy young elite athletes aged 11-15 years. *J. Strength Cond. Res.* 2014, *28*(5), 1328–34, doi: 10.1519/JSC.0000000000000280.

73. Bernard, J-C.C.; Bard, R.; Pujol, A.; Combey, A.; Boussard, D.; Begue, C.; et al. Muscle assessment in healthy teenagers. Comparison with teenagers with low back pain. *Ann. Readapt. Med. Phys.* 2008, *51*(4), 274–83, doi: 10.1016/j.annrmp.2008.03.010.

74. Ervin, R.; Fryar, C.; Wang, C-Y.; Miller, I.; Ogden, C. Strength and body weight in US children and adolescents. *Pediatrics.* 2014, *134*(3), e782–9, doi: 10.1542/peds.2014-0794.

75. Mac-Thiong, J.M.; Berthonnaud, É.; Dimar, J.R.I.; Betz, R.R.; Labelle, H. Sagittal Alignment of the Spine and Pelvis During Growth. *Spine (Phila Pa 1976).* 2004, *29*(15), 1642–7, doi: 10.1097/01.BRS.0000132312.78469.7B.

76. Widhe, T. Spine: posture, mobility and pain. A longitudinal study from childhood to adolescence. *Eur. Spine J.* 2001, *10*(2), 118–23, doi: 10.1007/s005860000230.

77. Poussa, M.S.; Heliövaara, M.M.; Seitsamo, J.T.; Könönen, M.H.; Hurmerinta,

K.A.; Nissinen, M.J. Development of spinal posture in a cohort of children from the age of 11 to 22 years. *Eur. Spine J.* 2005, *14*(8), 738–42, doi: 10.1007/s00586-004-0701-9.

78. Dejanovic, A.; Harvey, E.; Andersen, J.; Mcgill, SM. Do Anthropometric Measures Influence Torso Muscle Endurance Profiles of Children Aged 7 to 14? *Adv. Phys. Educ.* 2012, *2*(4), 187–96, doi: 10.4236/ape.2012.24032.
79. Marras, W.S.; Jorgensen, M.J.; Granata, K.P.; Wiand, B. Female and male trunk geometry: Size and prediction of the spine loading trunk muscles derived from MRI. *Clin. Biomech.* 2001, *16*(1), 38–46, doi: 10.1016/S0268-0033(00)00046-2.
80. Mannion, A.F.; Connolly, B.; Wood, K.; Dolan, P. The use of surface EMG power spectral analysis in the evaluation of back muscle function. *J. Rehabil. Res. Dev.* 1997, *34*(4), 427–39.
81. Dos Santos, M.A.; Lunkes, L.C.; Ribeiro, A de O.; Souza, A de C. Low back pain and risk factors during the third infancy. *Fisioter. Mov.* 2019, *32*, e003210, doi: 10.1590/1980-5918.032.ao10.
82. Graup, S.; de Araújo Bergmann, M.L.; Bergmann, G.G. Prevalence of nonspecific lumbar pain and associated factors among adolescents in Uruguaiana, state of Rio Grande do Sul. *Rev. Bras. Ortop.* 2014, *49*(6), 661–7, doi: 10.1016/j.rboe.2014.10.003.
83. Minghelli, B.; Oliveira, R.; Nunes, C. Non-specific low back pain in adolescents from the south of Portugal: prevalence and associated factors. *J. Orthop. Sci.* 2014, *19*(6), 883–92, doi: 10.1007/s00776-014-0626-z.
84. O'Sullivan, P.B.; Smith, A.J.; Beales, D.; Straker, L.M. Association of Biopsychosocial Factors With Degree of Slump in Sitting Posture and Self-Report of Back Pain in Adolescents: A Cross-Sectional Study. *Phys. Ther.* 2011, *91*(4), 470–83, doi: 10.2522/ptj.20100160.
85. Santonja, F. *Exploración clínica y radiográfica del raquis sagital: sus correlaciones*. [Thesis]. Universidad de Murcia; 1993.
86. Cavalcante Filho, D.E.A.; de Lima Viana, C.N.; de Santana Cabral, M.P.S.; Cunha, F.V.M.; de Sousa Pacheco, F.; da Silva Brito, A.K.; et al. Low back pain in adolescents: A school screening. *J. Hum. Growth Dev.* 2014, *24*(3), 347–53, doi: 10.7322/jhdg.88975.
87. Noormohammadpour, P.; Borghei, A.; Mirzaei, S.; Mansournia, M.A.; Ghayour-Najafabadi, M.; Kordi, M.; et al. The Risk Factors of Low Back Pain in Female

High School Students. *Spine (Phila Pa 1976)*. 2018, *44*(6), E357–65, doi: 10.1097/BRS.0000000000002837.

88. Kujala, U.M.; Taimela, S.; Oksanen, A.; Salminen, J.J. Lumbar mobility and low back pain during adolescence - A longitudinal three-year follow-up study in athletes and controls. *Am. J. Sports Med.* 1997, *25*(3), 363–8.
89. Salminen, J.J.; Maki, P.; Oksanen, A.; Pentti, J. Spinal mobility and trunk muscle strength in 15-year-old schoolchildren with and without low-back pain. *Spine (Phila Pa 1976)*. 1992, *17*(4), 405–11.
90. Feldman, D.; Shrier, I.; Rossignol, M.; Abenhaim, L.; Ehrmann-Feldman, D.; Rossignol, M.; et al. Risk factors for the development of low back pain in adolescence. *Am. J. Epidemiol.* 2001, *154*(1), 30–6, doi: 10.1093/aje/154.1.30.
91. Jones, M.A.; Stratton, G.; Reilly, T.; Unnithan, V.B. Biological risk indicators for recurrent non-specific low back pain in adolescents. *Br. J. Sports Med.* 2005, *39*(3), 137–40, doi: 10.1136/bjsm.2003.009951.
92. Sjolie, A.N. Low-back pain in adolescents is associated with poor hip mobility and high body mass index. *Scand. J. Med. Sci. Sports*. 2004, *14*(3), 168–75, doi: 10.1046/j.1600-0838.2003.00334.x.
93. Harreby, M.; Nygaard, B.; Jessen, T.; Larsen, E.; Storr-Paulsen, A.; Lindahl, A.; et al. Risk factors for low back pain in a cohort of 1389 Danish school children: An epidemiologic study. *Eur. Spine J.* 1999, *8*(6), 444–50, doi: 10.1007/s005860050203.
94. González-Gálvez, N.; Vaquero-Cristobal, R.; López-Vivancos, A.; Albaladejo-Saura, M.; Marcos-Pardo, P.J. Back Pain Related with Age, Anthropometric Variables, Sagittal Spinal Curvatures, Hamstring Extensibility, Physical Activity and Health Related Quality of Life in Male and Female High School Students. *Int. J. Environ. Res. Public Health*. 2020, *17*(7293), 1–16, doi: 10.3390/ijerph17197293.
95. Perry, M.; Straker, L.M.; O'Sullivan, P.; Smith, A.; Hands, B.; M., P.; et al. Fitness, Motor Competence, and Body Composition Are Weakly Associated With Adolescent Back Pain. *J. Orthop. Sport Phys. Ther.* 2009, *39*(6), 439–49, doi: 10.2519/jospt.2009.3011.
96. Smith, A.; Beales, D.; Sullivan, P.O.; Bear, N.; Straker, L. Low back pain with impact at 17 years of age is predicted by early adolescent risk factors from multiple domains: analysis of the Western Australian pregnancy cohort (raine)

study. *J. Orthop. Sports Phys. Ther.* 2017, *47*(10), 752–62, doi: 10.2519/jospt.2017.7464.

97. Smith, A.J.; O'Sullivan, P.B.; Campbell, A.; Straker, L. The relationship between back muscle endurance and physical, lifestyle, and psychological factors in adolescents. *J. Orthop. Sports Phys. Ther.* 2010, *40*(8), 517–23, doi: 10.2519/jospt.2010.3369.
98. Hebert, J.J.; Leboeuf-Yde, C.; Franz, C.; Lardon, A.; Hestbæk, L.; Manson, N.; et al. Pubertal development and growth are prospectively associated with spinal pain in young people (CHAMPS study-DK). *Eur. Spine J.* 2019, *28*(7), 1565–71, doi: 10.1007/s00586-019-05905-6.
99. Evans, K.; Refshauge, K.M.; Adams, R.; Aliprandi, L. Predictors of low back pain in young elite golfers: A preliminary study. *Phys. Ther. Sport.* 2005, *6*(3), 122–30, doi: 10.1016/j.ptsp.2005.05.003.
100. Cejudo, A.; Ginés-Díaz, A.; Rodríguez-Ferrán, O.; Santonja-Medina, F.; Sainz de Baranda, P. Trunk Lateral Flexor Endurance and Body Fat: Predictive Risk Factors for Low Back Pain in Child Equestrian Athletes. *Children.* 2020, *7*(10), 172, doi: 10.3390/children7100172.
101. Correia, J.P.; Oliveira, R.; Vaz, J.R.; Silva, L.; Pezarat-Correia, P. Trunk muscle activation, fatigue and low back pain in tennis players. *J. Sci. Med. Sport.* 2016, *19*(4), 311–6, doi: 10.1016/j.jsams.2015.04.002.
102. Mohseni-Bandpei, M.A.; Bagheri-Nesami, M.; Shayesteh-Azar, M. Nonspecific low back pain in 5000 Iranian school-age children. *J. Pediatr. Orthop.* 2007, *27*(2), 126–9, doi: 10.1097/BPO.0b013e3180317a35.

ESTUDIO 3. A Meta-Analysis of the Reliability of Four Field-Based Trunk Extension Endurance Tests

3.1. Introduction

In the last two decades, there has been an increased interest in the assessment of trunk extensor muscle endurance, as deficits in trunk extensor endurance and imbalances between trunk muscle groups have been suggested as being primary risk factors for low back disorders [1–4] and could negatively affect sports performance [5,6].

Although several sophisticated laboratory-based tests have been developed to quantify trunk extensor muscle endurance (i.e., force platforms, isokinetic dynamometers) [7–9], field-based tests seem to be the most popular tests, probably because they are portable, cost-effective, easy to use, and time-efficient methods [10]. Different trunk extension endurance field-based tests have been widely used, including (a) isometric endurance tests (i.e., Biering-Sorensen test [2], prone isometric chest raise test [11], and prone double straight-leg raise test [12]), which involve maintaining a position against gravity for as long as possible, and (b) dynamic endurance tests (i.e., dynamic extensor endurance test [3,13]), which consist of performing as many repetitions as possible in a given time or with a certain cadence until exhaustion. However, the validity and reliability of their outcomes must be determined [14] before these field-based tests can be used to identify deficits in trunk extensor muscle endurance, in order to explore imbalances between trunk muscle groups and to establish progress from training and/or rehabilitation programs. The four trunk extension endurance measures identified have been considered operationally valid by medical (American College of Sports Medicine [15]), sport (Swiss Olympic Medical Centers [16]), and educational (Cooper Institute [17]) organizations. Furthermore, their measures have been shown to be sensitive enough to detect trunk extensor muscle endurance deficits in patients with chronic low back pain [18,19], and are included in many prominent sports medicine textbooks [15–17].

Two different types of reliability should be considered before choosing an appropriate test for research and clinical purposes [20,21]: inter-tester reliability and intra-tester reliability. Inter-tester reliability provides information regarding the degree to which measures taken by different testers using identical test protocols on the same cohort of individuals are similar or consistent [20]. Intra-tester reliability provides information regarding the degree to which several measurements taken at different times for the same test by the same tester are similar [22]. The intra-tester reliability can be determined using short (generally within a day: intra-session) or long (generally more than one day: inter-session or test–retest) time intervals to separate the testing sessions [14,20].

In the scientific literature, several studies have examined the inter- [18,23,24] and intra-tester (intra- [18,25,26] and inter-session [9,11,27]) reliability of the four field-based trunk extension endurance tests in different cohorts of individuals (e.g., athletes, children, adolescents, sedentary adults). The findings of these studies, and for both types of reliability, have shown a large degree of heterogeneity, with intraclass correlation coefficient (ICC) scores ranging from 0.77 to 0.99 and from 0.20 to 0.99 for the inter- and intra-tester reliability, respectively.

The large degree of heterogeneity presented in the ICC scores reported suggests that the reliability outcomes of a physical performance measure could be influenced by several factors (the tester's experience or training in administering the test, variations in the assessment methodology, and participant-related variability, etc.). If practitioners do not consider these factors, they might select a trunk extension endurance test with an inter- and intra-tester reliability that is inappropriate for their given population. In practical terms, the selection of an inappropriate test can lead not only to an inaccurate diagnosis of normality or deficits but also an incorrect evaluation of the effectiveness of a training or rehabilitation intervention for improving or maintaining trunk extensor endurance. Frequently, studies exploring trunk extension endurance do not report the reliability estimates of the test they are using with the given population [28–30], or the reliability coefficients obtained in previous studies are simply cited [3,31,32]. This characteristic of referencing the reliability coefficients from prior studies instead of reporting the reliability obtained with their data has been coined reliability induction (RI) [33].

The reliability generalization (RG) is a meta-analytical approach that emerges as a criticism of the widespread practice of RI. The purpose of this method is to estimate the average reliability of the scores of a given test, as well as to determine the variability of the reliability coefficients reported by the different studies that have used this test. Moreover, if the variability is very high, another aim is to explore which characteristics of the studies may be statistically associated with the reliability estimates [34–36]. To the best of the authors' knowledge, there are no RG meta-analyses published to date concerning trunk extension endurance measures obtained through field-based tests. This information might be useful for practitioners because it can aid them in deciding on the best field-based test/s (in term of reliability) to assess trunk extensor muscle endurance according to the characteristics of their patients or athletes.

Therefore, the main purpose of the current study was to conduct a RG meta-analysis (a) to obtain combined reliability estimates of trunk extension endurance measures obtained through four field-based tests, (b) to identify which characteristics of the studies may influence the variability of the reliability coefficients, and (c) to determine the RI practice in studies that have used trunk extensor muscle endurance tests.

3.2. Methods

The current RG meta-analysis (PROSPERO ID: CRD42019123179) was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [37]. The PRISMA checklist is presented in Appendix 1 in the Supplementary Materials section.

3.2.1. Study Selection

In this RG meta-analysis, the following inclusion criteria were considered: (a) being an empirical research study (psychometric study) in which the original and modified versions of the following field-based tests were applied: Biering-Sorensen test [2], prone isometric chest raise test [11], prone double straight-leg raise test [12], and dynamic extensor endurance test [3] (a detailed description of each test is presented in Table 19); (b) being written in English or Spanish; (c) being published before March 2019; (d) using a sample of at least 10 participants; and (e) reporting the ICC as a measure of reliability (for both the inter- and intra-tester (intra-session or internal consistency and inter-session or test-retest reliability) reliability) of any of the aforementioned field-based tests or providing sufficient data from which this coefficient could be calculated through standardized equations. The selection of the ICC as a valid measure of reliability was made following the consensus-based standards for the selection of health measurement instruments (COSMIN) statement [38–40]. Studies whose main objective was not the analysis of the reliability of the trunk extension endurance measures obtained through any of the four aforementioned field-based tests, but which reported reliability data were also eligible for inclusion. The same inclusion criteria were considered for selecting studies that induced reliability, except (d) and (e).

Literature reviews, abstracts, editorial commentaries, and letters to the editor were excluded. Finally, some authors were contacted to provide missing data or to clarify if data were duplicated in other publications. Incomplete data or data from an already included study were also excluded.

Table 19. Trunk extension endurance field-based tests description.

Isometric Endurance Field-Based Tests	
Biering-Sorensen test	
	The test consists of assessing how many seconds the participant can keep the unsupported upper part of the body (from the upper border of the iliac crest) horizontal while placed prone with the buttocks and legs fixed to the table bench by three wide canvas straps, with the arms across the chest. The test is continued until the participant could no longer control his/her posture for a maximum of 240 s.
Prone isometric chest raise test	
	The test consists of assessing how many seconds the participant can keep the sternum off the floor while placed prone with the arms along the body. A small pillow is placed under the iliac crest to decrease the lumbar lordosis. The subject is asked to maintain the position for as long as possible, not exceeding a 5 min time limit.
Prone double straight-leg raise test	
	The test consists of assessing how many seconds the participant can keep both legs raised with the knees off the mat while placed prone with hips extended, the hands underneath the forehead and the arms perpendicular to the body. The test is continued until the participant can no longer maintain knee clearance.
Dynamic Endurance Field-Based Test	
Dynamic extensor endurance test	
	The test consists of assessing how many repetitions the participant can perform while placed prone with the unsupported upper part of the body (from the upper border of the iliac crest). The arms are positioned along the body and the buttocks and legs are fixed by three straps. With the spine kept straight, the subject is instructed to extend the trunk to neutral and then to lower the upper body 45 degrees. A repeated beat guided the subject to maintain a cadence of 25 repetitions per minute until exhaustion.

3.2.2. Search Strategy

A systematic computerized search was conducted up to 28th February 2019 in the databases MEDLINE, PubMed, Web of Science, and Scopus. In addition, a supplementary search in Google Scholar was also performed. Relevant keywords were used to construct Boolean search strategies, including terms such as reliability, extensor muscles, trunk, core, endurance, performance, and test (see Appendix 2). Then, the reference lists of included studies were screened for potentially eligible studies. Two reviewers independently (M.T.M. and P.S.B.) selected studies for inclusion in a two-step process. First, studies were screened on the basis of title and abstract. In a second stage, full-text studies were reviewed to identify those studies that met the eligibility criteria. Disagreements were resolved through consensus or by consulting a third reviewer (F.A.).

3.2.3. Data Extraction

To guarantee the maximum possible objectivity, a codebook was produced that specified the standards followed in coding each of the characteristics of the studies selected (see Appendix 3). The moderator variables of the eligible studies were coded and grouped into four categories: (1) general study descriptors (authors, publication year, country, and study objective), (2) description of the study population (sample size, age, sex distribution, and target population), (3) description of the field-based test (version, protocol, and scores), and (4) type of reliability analysed (inter-tester and/or intra-tester) and main characteristics of the study design (familiarization sessions, number of measurements, number of testers, time interval between measurements, test conditions, etc.).

3.2.4. Reliability Estimates

In this RG study, the ICC was extracted from the eligible studies as the only valid measure of inter- and intra-tester reliability. There are other coefficients such as the standard error of measure (SEM), Pearson correlation, and minimal detectable change with a 95% confidence level (MDC95) that have also been proposed as measures of reliability [20,41]. However, the scarce number of references that reported these types of reliability coefficients [42–44] did not allow for any separate meta-analyses to be conducted.

In order to maximize the number of studies included in this RG, when an article only reported the SEM, ICC was estimated by

$$ICC = \sqrt{1 - \frac{SEM^2}{SD^2}}, \quad (1)$$

with SD being the standard deviation of the test scores.

Likewise, when an article only reported the MDC, it was transformed into SEM using the following formula:

$$SEM = \frac{MDC}{1.96 \times \sqrt{2}} \quad (2)$$

Later, the SEM was converted into ICC using the previously stated formula.

3.2.5. Quality Assessment

The COSMIN methodology was used to evaluate the quality of eligible studies for the RG meta-analysis, which consists of three sub-steps [38–40]. First, the risk of bias in each study was assessed using the COSMIN risk of bias check-list [39]. The check-list contains standards referring to design requirements and preferred statistical methods of studies on measurement properties. For each measurement property, a COSMIN box was developed, containing all standards needed to assess the quality of a study on that specific measurement property. Specifically, Box 6 regarding reliability was used for this RG meta-analysis. Each standard of the box was rated as very good, adequate, doubtful, or inadequate quality. To determine the overall rating of the quality of every single study, the lowest rating of any standard in the box was taken (i.e., "the worst score counts" principle) [45]. Second, from the data extracted on the description of the study population and the results on the reliability coefficients, the result per study was rated against the criteria for good measurement properties [38] as sufficient (+ (ICC $\geq$ 0.70)), insufficient (− (ICC $<$ 0.70)), or indeterminate (? (ICC not reported)). Finally, the results from different studies on one measurement property were statistically pooled in a meta-analysis and the quality of the evidence was graded (high, moderate, low, or very low evidence) using the modified Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. Four of the five GRADE factors were used in this meta-analysis: risk of bias (i.e., the methodological quality of the studies), inconsistency (i.e., unexplained inconsistency of results across studies), imprecision (i.e., total sample size of the available studies), and indirectness (i.e., evidence from different populations than the population of interest in the review). The fifth factor, publication bias, is difficult to assess in studies on measurement properties due to a lack of registries for these types of studies. Therefore, we did not take this factor into account in this meta-analysis [38–40]. The starting point is always the assumption that the pooled or overall result is of high quality. The quality of evidence is subsequently downgraded by one or two levels per factor to moderate, low, or very low when there is a risk of bias, inconsistency, imprecision, or indirect results [38]. Appendix 4 displays a brief description of each step of the COSMIN methodology.

The data extraction and quality assessment were double-coded (M.T.M. and P.S.B.) to assess the inter-coder reliability of the coding process. Two authors working independently randomly coded 50% of the studies. For the quantitative moderator variables, ICC3,1 were calculated, whereas for the qualitative moderator variables, we applied Cohen's kappa coefficients. In general, the agreement coefficients between the two authors were satisfactory as proposed by Orwin and Vevea [46], with the kappa coefficients ranging between 0.74 and 1, and the ICC ranging between 0.96 and 1. Inconsistencies between the two coders were resolved by consensus, and when these were due to ambiguity in the coding book, this was corrected. As before, any disagreement was resolved by mutual consent or in consultation with a third reviewer (FA).

3.2.6. Data Synthesis and Analysis

Separate meta-analyses were carried out for the different field-based tests and types of reliability analysis to avoid dependence problems, given that each study could report more than one reliability coefficient.

Statistical analyses were completed by assuming a random-effects model, and each reliability coefficient was weighted by the inverse variance [47]. Reliability coefficients were transformed into Fisher's Z to normalize their distribution and to stabilize the variances [36]. For each meta-analysis, an average reliability coefficient (ICC+) and a 95% confidence interval (95% CI) were calculated [48].

The heterogeneity exhibited by the reliability coefficient, which represents the percentage of total variation across all studies due to between-study heterogeneity, was assessed by constructing a forest plot and by calculating the Q statistic and the I2 index. Analyses of potential moderator variables were carried out when the Q statistic was statistically significant, the I2 index was over 25%, and there were at least 20 reliability coefficients [49]. The only test that fulfilled these conditions was the Biering-Sorensen test for intra-tester reliability ($k = 28$ studies).

The influence of qualitative moderator variables (i.e., sample type (children and adolescents, adults), target population (community, clinical), physical activity level (sedentary, active), test duration (until exhaustion, until 240 s), among others) on the reliability coefficients were explored through analysis of variance (ANOVAs) and assuming mixed-effects model. Meta-regressions were also applied to test the influence of continuous moderators on the reliability coefficients assuming mixed-effects model, such as final sample size, average percentage of females, average age, percentage of attrition, number of measurements, number of testers, and time interval between measurements. QB and Qw for ANOVAs and QR and QE statistics for meta-regressions were calculated to test the statistical significance of each moderator variable and to assess the model misspecification, respectively [50]. In addition, the proportion of variance accounted for by the moderator variables was estimated with R2 following Raudenbush's proposal [51]. In order to find the subset of moderator variables that can explain most of the reliability coefficient variability, a multiple meta-regression model (by assuming a mixed-effects model) was adjusted.

To facilitate the interpretation of the results, the average reliability coefficients and their confidence limits were back-transformed to the original metric of reliability coefficients. Furthermore, it has been suggested that for studies conducted in the sports medicine field of knowledge and aimed at analysing the reliability (inter- and intra-tester) of quantitative physical performance measures, ICC values of 0.8 to 0.9 may be considered as acceptable, but values higher than 0.9 are desirable [14].

The statistical analyses were carried out with the software Comprehensive Meta-analysis 3.3 (BioStat, Englewood, New Jersey, United States) [52].

3.3. Results

3.3.1. Study Selection

A total of 1452 references were identified with all search strategies, from which 494 were excluded in the first screening as duplicates (34%). A total of 674 studies (46.4%) were eliminated after reading the title and abstract. Another 181 studies (12.4%) were removed after reading the full-text—120 studies did not use any trunk extensor muscle endurance field-based test (8.2%) and 61 articles did not meet the established inclusion criteria (4.2%).[53–69].

This search process identified 103 empirical studies that met the inclusion criteria (7.1%), in which 28 articles (1.9%) reported some reliability coefficients [8–11,18,23–27,43,53–69] (resulting in 43 cohort groups, as 9 studies had more than 1 group) and 75 articles (5.1%) induced the reliability. Figure 15 shows the flow chart of the selection process of the studies.

3.3.2. Descriptive Characteristics of the Selected Studies for the RG Meta-Analysis

Appendix 5 provides a descriptive summary of the characteristics of the included studies for the RG meta-analysis. The studies selected were carried out between 1986 and 2018 and comprised participants from four continents (Europe [8,10,24,43,53–57,59–61,63,65,66], America [9,25–27,64,67–69], Asia [11,18,58,62], and Oceania [23]). Only one study was written in Spanish [60], and the rest of the selected studies were written in English [8–11,18,23–27,43,53–59,61–69].

The total sample size was 1097 participants, with an average of 25.5 subjects per cohort group (minimum = 10 [43,53,59,62,63] and maximum = 100 [11]). Three studies used healthy children [57] and adolescents [27,62]. Furthermore, 14 studies employed asymptomatic samples [9,10,18,24,25,27,53,56,57,59,60,62,63,68,69], 8 used patients with low back pain (LBP) [8,54,55,58,64–67], and 6 included both types of samples [11,23,26,43,61].

Concerning the type of field-based test used, 25 studies (89.2%) investigated the reliability of the back extensor endurance measures obtained through different modified versions of the Biering-Sorensen test [8–10,18,23–27,43,53,54,56,57,59–69], 4 studies (14.3%) used the prone isometric chest raise test (two of them used the original version [11,58] and another two used a modified version [18,55]), 2 studies employed a modified version of the dynamic extensor endurance test [27,62], (7.1%) and 1 study used the original prone double straight-leg raise test [18] (3.6%).

The most common types of reliability coefficients reported in the selected articles were the intra-tester and inter-session reliability (24 studies [8–11,24,26,27,43,53–66,68,69] that reported 37 coefficients), followed by the intra-tester and intra-session reliability (5 studies [18,25,26,43,67] that reported 16 coefficients), and the inter-tester reliability with 10 coefficients reported from 5 different studies [18,23,24,26,66].

3.3.3. Quality of the Selected Studies for the RG Meta-Analysis

Regarding the use of the Biering-Sorensen test in the studies, the COSMIN risk of bias checklist reported very good methodological quality scores for seven studies (five studies that evaluated the intra-tester and the inter-session reliability [8,10,59,61,66] and two studies for the inter-tester reliability [24,66]) and adequate quality scores in one study that analysed the intra-tester and inter-session reliability [26]. In contrast, 13 studies [9,24,27,43,53,54,56,57,60,64,65,68,69] that analysed the intra-tester and inter-session reliability obtained a doubtful methodological quality, as well as all of the studies that assessed the intra-tester and intra-session reliability [18,25,26,43,67] and three studies that evaluated the inter-tester reliability [18,23,26]. Two studies that analysed the intra-tester and inter-session reliability showed inadequate methodological quality [62,63].

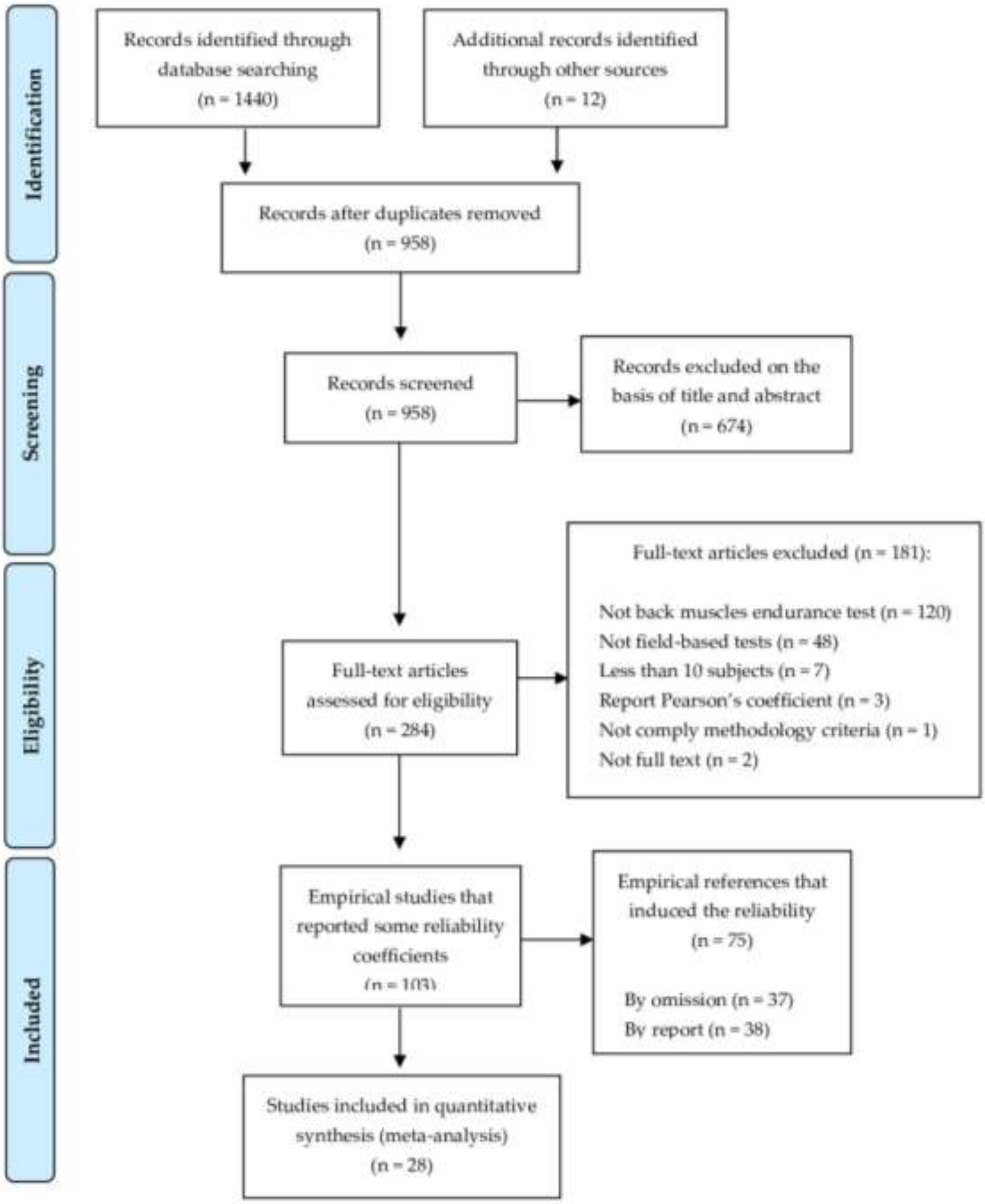


Figure 15. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the literature search.

Concerning the use of the prone isometric chest raise test, three studies presented doubtful methodological quality scores for intra-tester and inter-session reliability [11,55,58], the same as for intra-tester and intra-session and inter-tester reliability where only one study was found [18]. It was the same with the prone double straight-leg test, which reported a doubtful quality for the only study found [18]. Finally, the two studies that used the dynamic extensor endurance test exhibited doubtful [27] and inadequate [62] methodological qualities for the intra-tester and inter-session reliability.

The detailed data of the COSMIN risk of bias check-list and criteria for good measurement properties are presented in Appendix 6. Likewise, a summary of findings (SoF) per type of reliability and field-based test, including the pooled results of the measurement properties, the overall rating based on the inconsistency of results (i.e., sufficient (+), insufficient (−), or indeterminate (?)), and the grading of the quality of evidence (i.e., high, moderate, low, very low) is presented in Table 20. In general, the Biering-Sorensen test is the only test that presented moderate quality of evidence for inter-tester and intra-tester (inter-session) reliability.

Table 20. Summary of findings (SoF).

Reliability	Pooled Result [a]	Overall Rating [b]	Quality of Evidence [c]
Biering-Sorensen Test			
Inter-tester reliability	ICC = 094 (0.84–0.98) Consistent results Sample size = 215	Sufficient	Moderate (as there are multiple doubtful [18,23,27] and two very good studies [64,69])
Intra-tester (intra-session) reliability	ICC = 0.88 (0.83–0.92) Consistent results Sample size = 258	Sufficient	Low (as all studies are doubtful [18,25,27,44,70])
Intra-tester (inter-session) reliability	ICC = 0.88 (0.80–0.92) Consistent results Sample size = 688	Sufficient	Moderate (as there are multiple doubtful and five very good studies [8,10,61,63,69])
Prone Isometric Chest Raise Test			
Inter-tester reliability	ICC = 0.90 (0.80–0.95) Sample size = 30	Indeterminate	Did not pool the results or grade the evidence due to there being one study available [18]
Intra-tester (intra-session) reliability	ICC = 0.90 (0.83–0.94) Sample size = 30	Indeterminate	Did not pool the results or grade the evidence due to there being one study available [18]
Intra-tester (inter-session) reliability	ICC = 0.95 (0.91–0.97) Consistent results Sample size = 236	Sufficient	Low (as all studies are doubtful [11,56,60])
Prone Double Straight-Leg Test			
Inter-tester reliability	ICC = 0.83 (0.67–0.93) Sample size = 30	Indeterminate	Did not pool the results or grade the evidence due to there being one study available [18]
Intra-tester (intra-session) reliability	ICC = 0.86 (0.77–0.92) Sample size = 30	Indeterminate	Did not pool the results or grade the evidence due to there being one study available [18]
Dynamic Extensor Endurance Test			
Intra-tester (inter-session) reliability	ICC = 0.99 (0.88–1.00) Consistent results Sample size: 82	Sufficient	Low (as there is one inadequate study [65])

[a] Pooled results obtained from mean reliability analysis and adjusted according to the publication bias analysis. [b] Overall rating was graded as sufficient (intraclass correlation coefficient (ICC) > 0.70), insufficient (ICC < 0.70), or indeterminate (either ICC reported by just one study or by none). [c] Quality of evidence (high, moderate, low, very low) based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach that uses three factors: (1) risk of bias assessed with the consensus-based standards for the selection of health measurement instruments (COSMIN) risk of bias check-list; (2) inconsistency, solved by pooling the results; and (3) imprecision, the total sample included in the studies. When the total sample size of the pooled studies is below 100, downgrade with one level and with two levels when the total sample size is below 50. ICC: intraclass correlation coefficient.

3.3.4. Effect Sizes

Primary Outcomes

Table 21 presents the average reliability of the trunk extensor measures obtained from the different field-based tests and separately by type of reliability.

Table 21. Average reliability, 95% confidence intervals, and heterogeneity statistics for each field-based test and reliability type.

	k	ICC_+	ICC_L	ICC_U	*Q*	*df*	*p*	I^2
Biering-Sorensen Test:								
▪ Inter-tester reliability	8	0.94	0.84	0.98	102.37	7	<0.001	93.1%
▪ Intra-tester and intra-session reliability	12	0.88	0.83	0.92	22.69	11	0.02	51.2%
▪ Intra-tester and inter-session reliability	27	0.88	0.80	0.92	234.40	26	<0.001	88.9%
Prone Isometric Chest Raise test:								
▪ Inter-tester reliability	1	0.90	0.80	0.95	--	--	--	--
▪ Intra-tester and intra-session reliability	2	0.90	0.83	0.94	0.034	1	0.853	0%
▪ Intra-tester and inter-session reliability	5	0.95	0.91	0.97	13.57	4	0.009	70.5%
Prone Double Straight-Leg Test:								
▪ Inter-tester reliability	1	0.83	0.67	0.92	--	--	--	--
▪ Intra-tester and intra-session reliability	2	0.86	0.77	0.92	0.08	9	0.777	0%
Dynamic Extensor Endurance Test:								
▪ Intra-tester and inter-session reliability	5	0.99	0.88	1.00	15.73	4	0.003	74.5%

k: number of cohorts, ICC+: mean intraclass correlation coefficient, ICCL and ICCU: 95% lower and upper CI for ICC+, Q: heterogeneity statistic, DF: degrees of freedom for Q statistic, p: probability level associated to Q statistic, I2: heterogeneity index.

Figure 16 displays a summary of the intra-tester and inter-session reliability obtained in the studies that applied the Biering-Sorensen test.

In the different meta-analyses carried out, the effect sizes exhibited a moderate to large heterogeneity (based on the Q statistics and the I2 indices), supporting the decision of applying the random-effects model.

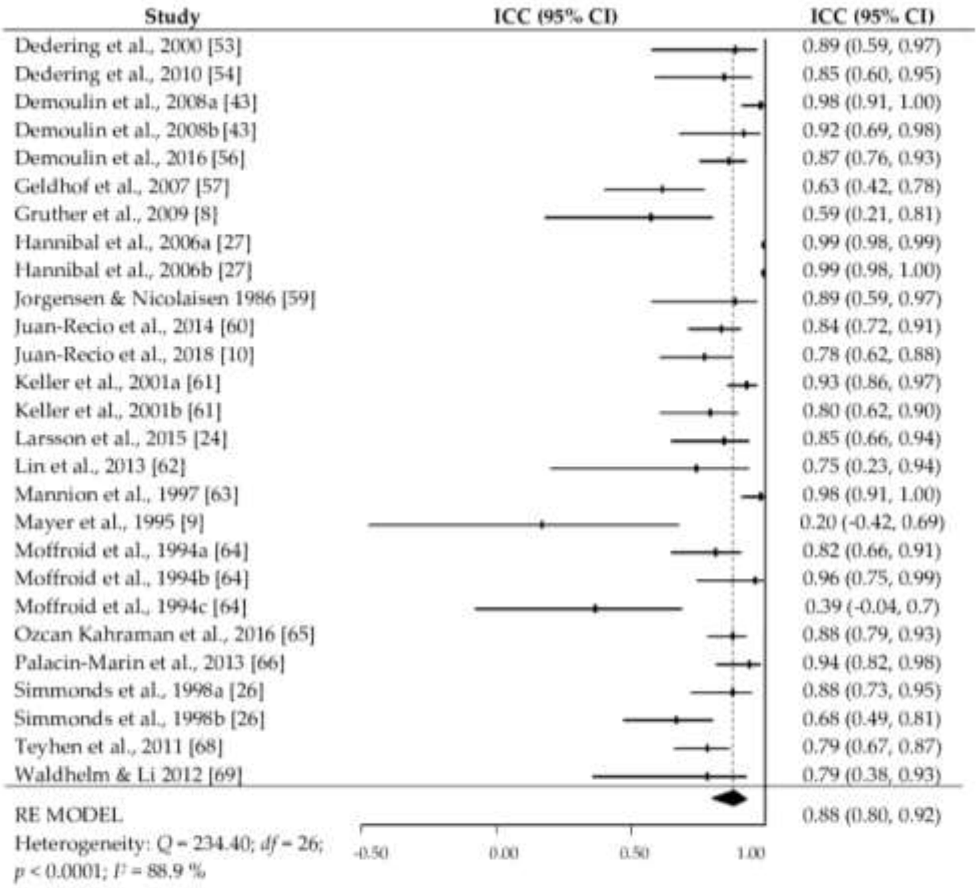


Figure 16. Forest plot of the intra-tester and inter-session reliability obtained in the studies that applied the Biering-Sorensen test.

<u>Analysis of the Moderator Variables</u>

As only the intra-tester and inter-session reliability for the Biering-Sorensen test presented more than 20 reliability estimates, the analyses of moderator variables were carried out exclusively for this field-based test and type of reliability, in order to examine the influence of qualitative and continuous moderator variables on the reliability coefficients.

Table 22 shows the results of the meta-regression analyses for each continuous moderator variable. None of the continuous moderators exhibited a statistically significant relationship with the intra-tester and the inter-session reliability of the Biering-Sorensen test.

Table 22. Results of the mixed-effects meta-regressions for the continuous moderator variables on the intraclass correlation coefficient (ICC) estimates obtained from intra-tester and inter-session reliability of the Biering-Sorensen test.

Moderator Variable	k	b_j	Q_R	p	Q_E	R^2
Publication year	27	0.003	0.03	0.855	232.21 *	0
Final sample size	27	−0.003	0.11	0.741	230.26 *	0
Sex (% female)	24	0.002	0.37	0.542	200.78 *	0
Mean age (years)	26	−0.017	1.62	0.203	210.81 *	0.01
SD age	19	−0.009	0.31	0.578	177.94 *	0
% attrition	27	−0.008	1.39	0.238	227.47 *	0
Number of measurements	27	−0.034	0.01	0.933	232.93 *	0
Time interval between measurement	26	−0.006	0.18	0.668	229.87 *	0
Mean test score from total sample	26	0.002	0.98	0.321	210.60 *	0
SD test score from total sample	21	0.009	0.94	0.331	201.42 *	0.02
Mean test score from reliability sample	24	0.001	0.42	0.514	207.27 *	0
SD test score from reliability sample	20	0.012	1.54	0.214	180.06 *	0.10

k: number of studies, b_j: regression coefficient for the predictor variable, Q_R: statistic for testing the statistical significance of the predictor variable, p: probability level for the Q_R statistic, Q_E: statistic for testing the model misspecification, R^2: proportion of variance explained by the predictor variable, * $p < 0.001$.

Regarding the qualitative moderator variables, Table 23 shows the ANOVA results. Concerning the description of the field-based tests, significant statistical differences were obtained in position control instruments (p < 0.0001) with an explained variance of 69%. Specifically, the studies that used a plumb-line to control position exhibited, on average, greater reliability (ICC+ = 0.99), followed by the stadiometer (ICC+ = 0.89), the light sensor (ICC+ = 0.87), the inclinometer (ICC+ = 0.83), and the visual control (ICC+ = 0.81). Similarly, statistically significant differences were observed (p = 0.045) between the use of a roman chair (ICC+ = 0.94) or a test bench (ICC+ = 0.85), with an explained variance of 32%.

Table 23. Results of the mixed-effects ANOVAs for the qualitative moderator variables on the intraclass correlation coefficient (ICC) estimates obtained from intra-tester and inter-session reliability of the Biering-Sorensen test.

Moderator Variables	k	ICC_+	95% CI		ANOVA Results
			ICC_L	ICC_U	
Reliability analysis was done with the same sample:					
Yes	24	0.88	0.81	0.92	$Q_B(1) = 0.92$, $p = 0.338$; $R^2 = 0.01$
No	3	0.77	0.26	0.94	$Q_W(25) = 217.83$, $p < 0.0001$
Sex:					
Males	9	0.88	0.77	0.94	$Q_B(2) = 5.98$, $p = 0.050$; $R^2 = 0.27$
Females	2	0.97	0.89	0.99	$Q_W(23) = 156.41$, $p < 0.0001$
Males and females	15	0.83	0.72	0.89	
Sample type:					
Children and adolescents	4	0.95	0.86	0.98	$Q_B(1) = 3.73$, $p = 0.053$; $R^2 = 0.18$
Adults	23	0.85	0.77	0.90	$Q_W(25) = 185.35$, $p < 0.0001$
Target population:					
Asymptomatic	18	0.88	0.80	0.93	$Q_B(1) = 0.24$, $p = 0.625$; $R^2 = 0$
Clinical	9	0.85	0.69	0.93	$Q_W(25) = 231.60$, $p < 0.0001$
Physical activity level:					
Sedentary	12	0.84	0.71	0.92	$Q_B(1) = 0.78$, $p = 0.332$; $R^2 = 0.03$
Recreationally active	15	0.89	0.81	0.94	$Q_W(25) = 214.29$, $p < 0.0001$
Validated modification:					
Yes	17	0.86	0.76	0.92	$Q_B(1) = 0.24$, $p = 0.626$; $R^2 = 0$
No	10	0.89	0.78	0.95	$Q_W(25) = 222.31$, $p < 0.0001$
Tool:					
Test bench	22	0.85	0.77	0.90	$Q_B(1) = 4.03$, $p = 0.045$; $R^2 = 0.32$
Roman chair	5	0.94	0.86	0.98	$Q_W(25) = 158.11$, $p < 0.0001$
Hands position:					
Crossed on the chest	22	0.88	0.80	0.93	$Q_B(2) = 0.37$, $p = 0.828$; $R^2 = 0$
Along the body	3	0.87	0.55	0.97	$Q_W(24) = 217.66$, $p < 0.0001$
At the level of the ears	2	0.80	0.23	0.96	
Part of the body on the edge:					
Not reported	12	0.90	0.80	0.95	
ASIS	6	0.88	0.68	0.95	$Q_B(3) = 1.26$, $p = 0.738\ 4$; $R^2 = 0$
Upper border of the iliac crest	6	0.85	0.62	0.95	$Q_W(23) = 212.02$, $p < 0.0001$
Pubis	3	0.77	0.28	0.94	
Test duration:					
Until exhaustion	25	0.88	0.82	0.93	$Q_B(1) = 1.24$, $p = 0.264$; $R^2 = 0.01$
Until 240 s	2	0.72	0.08	0.94	$Q_W(25) = 212.46$, $p < 0.0001$
Position control instruments:					
Visual	14	0.81	0.73	0.87	
Inclinometer	5	0.83	0.67	0.91	$Q_B(4) = 34.49$, $p < 0.0001$;
Stadiometer	4	0.89	0.78	0.95	$R^2 = 0.69$
Light sensor	2	0.87	0.59	0.96	$Q_W(22) = 72.29$, $p < 0.0001$
Plumb-line	2	0.99	0.98	0.99	
Familiarization session:					$Q_B(1) = 12.09$, $p = 0.0005$;
Yes	6	0.96	0.92	0.98	$R^2 = 0.34$
No	20	0.82	0.73	0.88	$Q_W(25) = 151.63$, $p < 0.0001$
Test conditions:					
Similar conditions	18	0.88	0.80	0.93	$Q_B(1) = 0.27$, $p = 0.605$; $R^2 = 0$
Unclear conditions	9	0.85	0.68	0.93	$Q_W(25) = 222.08$, $p < 0.0001$
The profession of tester:					
Sports sciences	6	0.92	0.80	0.97	$Q_B(2) = 1.88$, $p = 0.389$; $R^2 = 0.02$
Physical therapy	17	0.86	0.76	0.92	$Q_W(23) = 205.35$, $p < 0.0001$
Medicine	3	0.77	0.30	0.94	

Table 23. Results of the mixed-effects ANOVAs for the qualitative moderator variables on the intraclass correlation coefficient (ICC) estimates obtained from intra-tester and inter-session reliability of the Biering-Sorensen test.

Moderator Variables	k	ICC_+	95% CI		ANOVA Results
			ICC_L	ICC_U	
Continent:					$Q_B(1) = 0.0004$, $p = 0.984$; $R^2 = 0$ $Q_W(24) = 226.96$, $p < 0.0001$
Europe	16	0.88	0.78	0.93	
America	10	0.88	0.74	0.94	
Study objective:					$Q_B(1) = 0.02$, $p = 0.877$; $R^2 = 0$ $Q_W(25) = 231.27$, $p < 0.0001$
Psychometric	20	0.88	0.79	0.93	
Not psychometric	7	0.86	0.67	0.95	
Conflict of interest:					$Q_B(1) = 0.02$, $p = 0.896$; $R^2 = 0$ $Q_W(25) = 232.67$, $p < 0.0001$
No conflict	8	0.87	0.71	0.94	
Not reported	19	0.88	0.79	0.93	
COSMIN Risk of Bias check-list:					$Q_B(3) = 0.81$, $p = 0.848$; $R^2 = 0$ $Q_W(23) = 223.47$, $p < 0.0001$
Very good	6	0.85	0.62	0.94	
Adequate	2	0.80	0.22	0.96	
Doubtful	17	0.85	0.79	0.93	
Inadequate	2	0.92	0.56	0.99	

k: number of cohorts, ASIS: anterosuperior iliac spine, ANOVA: analysis of variance, ICC_+: mean reliability coefficient, ICC_L and ICC_U: lower and upper 95% confidence limits for ICC_+, Q_B: statistic for testing the statistical significance of the predictor variable, Q_W: statistic for testing the model misspecification, R^2: proportion of variance explained by the predictor variable.

Concerning the characteristics of the study design, the presence of a familiarization session before data collection showed a statistically significant relationship with the reliability estimates (p=0.0005). Specifically, carrying out a familiarization session showed larger average reliability (ICC+ = 0.96) than when it was not performed (ICC+ = 0.82), with an explained variance of 34%.

Finally, the sex of the sample approached statistical significance on the average reliability (p = 0.05, R2 = 0.27), with the female samples presenting a better reliability (ICC+ = 0.97) than the males (ICC+ = 0.88) or males and females together (ICC+ = 0.83). Similarly, the sample type showed a tendency towards statistical significance (p = 0.053, R2 = 0.18). The studies that used children or adolescents as the sample showed, on average, higher reliability (ICC+ = 0.95) than those that tested adults (ICC+ = 0.85).

An Explanatory Model

As a further step, a multiple meta-regression was applied to identify the subset of moderator variables that can explain most of the reliability coefficient variability. The predictors included in the model were selected on the basis of the ANOVAs and meta-regression results previously conducted. Thus, out of the five moderator variables that presented a statistically significant relationship with the reliability coefficients (sex, sample type, tool, position control instruments, and familiarization session), a multiple regression model with two predictors was the model that best explained the variability of the reliability coefficients—the type of sample (0, adults vs. 1, children/adolescents) and the familiarization session (0, No vs. 1, Yes).

Table 24 presents the results. These two moderator variables reached statistical significance (QR (2) = 20.57, p < 0.001) with 51% of variance accounted for. As shown in Table 24, both moderator variables exhibited a statistically significant relationship with the reliability coefficients, once the influence of the other moderator was controlled.

Table 24. Results of the multiple meta-regression model assuming a mixed-effects model.

Source	b_j	SE	Z	p
Intercept	1.09	0.11	9.89	< 0.001
Sample type	0.65	0.27	2.39	0.017
Familiarization session	0.68	0.24	2.88	0.004
Full model:	Q_R (2) = 20.57, R^2 = 0.51 Q_E (24) = 113.67, p < 0.001			

b_j: unstandardized regression coefficient, SE: standard error of b_j, Z: statistic for testing the statistical significance of each moderator variable, Q_R: statistic for testing the global significance of the model, Q_E: statistic for testing the model misspecification, R^2: proportion of variance accounted for by the full model.

3.3.6. Reliability Induction

Out of the 103 empirical studies that reported some reliability coefficients, 75 studies induced reliability (37 by omission and 38 by the report from other studies), which implies a 72.8% of RI. In particular, of these 38 studies that induced reliability by the report from other studies (41 cohorts), 32 induced reliability of the Biering-Sorensen test, 4 the prone isometric chest raise test, 2 the supine bridge test, 1 the prone double straight-leg raise test, and 2 the reliability of the dynamic extensor endurance test.

3.4. Discussion

The main purpose of the current RG meta-analysis was to estimate both the inter- and intra-tester (intra- and inter-session) reliability of the trunk extension endurance measures obtained through four field-based tests (Biering-Sorensen, prone isometric chest raise, prone double straight-leg raise, and dynamic extensor endurance), as well as to identify those characteristics (qualitative and quantitative moderators) of the studies selected that might have a meaningful influence in the degree of heterogeneity showed by the pooled reliability coefficients (ICC scores). A secondary purpose was to determine the RI practice in studies that used at least one of the four trunk muscle endurance field-based tests investigated in this RG meta-analysis.

The systematic literature review carried out showed that very few studies have analysed the inter-tester reliability of the trunk extension endurance measures obtained through field-based tests [18,23,24,26,66]. Specifically, most of these studies focused on the analysis of the inter-tester reliability of the trunk extension endurance measure obtained through the Biering-Sorensen test [18,23,24,26,66] (eight cohorts); the number of studies that explored the inter-tester reliability of the endurance measures recorded from the prone isometric chest raise [18] (one cohort) and the prone double straight-leg [18] (one cohort) tests were very limited, and no studies used the dynamic extensor endurance test. In particular, the results of this RG meta-analysis showed appropriate inter-tester reliability values (ICC+ > 0.80) for the trunk extension endurance measures obtained through the Biering-Sorensen (ICC+ = 0.94), the prone chest raise (ICC+ = 0.90), and the prone double straight-leg (ICC+ = 0.83) tests. However, and due to the scarce evidence available (one cohort with a small sample size (n = 30) was found), the inter-tester reliability estimates of the measures obtained through the prone isometric chest raise and prone double straight-leg tests should be considered with a degree of caution.

On the other hand, five studies [18,25,26,43,67] analysed the intra-tester reliability of the trunk extension endurance measures, using short periods (intra-session, within a day). Most of these studies focused exclusively on the endurance measure obtained through the Biering-Sorensen test, which reported appropriate reliability results (ICC+ = 0.88). For the rest of the trunk extensor endurance field-based tests, only one study was found that assessed the intra-tester reliability in its intra-session modality for the prone isometric chest raise and the prone double straight-leg tests [18], reporting acceptable reliability results (ICC+ = 0.90 and 0.86, respectively). However, these intra-tester reliability results for the prone isometric chest raise and the prone double straight-leg tests were not supported by strong evidence (only 1 cohort of 30 participants [18]) and hence should be considered with a degree of caution. There appears to be no study focusing on the intra-session reliability of the measure obtained from the dynamic extensor endurance test.

Contrarily, there was a higher number of studies that analysed the intra-tester and inter-session reliability of the trunk extension endurance measures obtained through the field-based tests selected (except the endurance measures obtained from the prone double straight-leg test, as no study has analysed its inter-session reliability). Different authors [70,71] have suggested that the scarce number of intra-session reliability coefficients reported in the literature may be explained by the fact that field-based tests are often used in longitudinal randomized controlled trials that analyse the chronic effects of training interventions rather than in studies that were conducted using short periods to analyse the acute effects of specific interventions. Therefore, the higher number of studies that have analysed the inter-session reliability in contrast with those that have analysed the intra-session reliability of the endurance measures seems to be in line with the research designs and interests currently present in the scientific literature. All the trunk extension measures analysed reported appropriate inter-session reliability scores, with ICC values ranging from ICC+ = 0.88 (Biering-Sorensen test) to 0.99 (dynamic extensor endurance test).

On the whole, only the trunk extension endurance measure obtained through the Biering-Sorensen test may present sufficient scientific evidence in terms of inter- (ICC+ = 0.94, based on 8 cohorts) and intra-tester (intra-session (ICC+ = 0.88, based on 12 cohorts) and inter-session (ICC+ = 0.88, based on 27 cohorts)) reliability to justify its use for practical and clinical purposes.

The analysis of the quantitative and qualitative factors or moderators that may have had an influence in the degree of heterogeneity found for the pooled reliability coefficients of the trunk extension endurance measures could only be carried out in the measurements obtained through the Biering-Sorensen test and for the inter-session subtype of intra-tester reliability because it was the only one with at least 20 reported reliability estimates. Thus, the results derived from the meta-regressions demonstrated that none of the factors/objects of the study had a statistically significant impact ($p > 0.05$) on the heterogeneity exhibited by the pooled inter-session reliability coefficient of the endurance measure obtained through the Biering-Sorensen test. It should be highlighted that even though a non-statistically significant relationship was found between the moderator "SD test score from reliability sample" and the inter-session reliability coefficient of the Biering-Sorensen test, it explained 10% of the variance. This finding agrees with the psychometric theory that states that the larger the SD of test scores, the greater the reliability that might be obtained [72].

Regarding the qualitative moderator analysis (ANOVAs), the instrument or tool used to control the participant’s position during the test (no instrument (visually controlled) vs. inclinometer vs. stadiometer vs. light sensor vs. plumb-line), the equipment used to perform the test (test bench vs. roman chair) and the presence (or not) of a familiarization session were the moderators that showed statistically significant effects ($p < 0.05$) on the inter-session reliability estimate of the endurance measure obtained through the Biering-Sorensen test.

During the execution of the Biering-Sorensen test, small movements sometimes occur naturally while participants are trying to maintain the required isometric horizontal position. These little movements may make the tester falsely assume a perception of fatigue in the participant who is being tested [23]. To minimize this source of error, it has been recommended that the tester should control the participant's position during the test through the use of a static visual reference. The results of this RG meta-analysis showed that the use of a plumb-line to control the participant's position (probably because, and unlike the other tools that only offer visual feedback of the participant's position to the tester, it also provides kinaesthetic feedback in the low back region to the participant who is being tested about the position that has to be maintained during the test) was the instrument that reported the highest inter-session reliability (ICC+ = 0.99), followed by the stadiometer (ICC+ = 0.89), light sensor (ICC+ = 0.87), inclinometer (ICC+ = 0.83), and visual control (ICC+ = 0.81).

The tool or equipment used to perform the Biering-Sorensen test also showed significant effects on the inter-session reliability estimate. Those studies that used the roman chair to perform the Biering-Sorensen test reported better inter-session reliability scores (ICC+ = 0.94) than those using the test bench (ICC+ = 0.85). However, this difference might be partially attributed to the imbalance distribution existing regarding the number of cohorts selected to carry out the inter-session reliability analysis of both sub-categories of this moderator. Thus, whereas five cohorts were used to calculate the inter-session reliability estimate of the trunk extensor endurance measure obtained through the Biering-Sorensen test using a roman chair, 22 cohorts were used to conduct the same analysis using a test bench.

One of the main sources of error in reliability studies is the presence of learning effects [14]. This effect was identified, as in studies in which participants were asked to carry out a familiarization session for the Biering-Sorensen test before the data collection reported statistically significantly better inter-session reliability coefficients (ICC+ = 0.96) than those that did not employ a familiarization session (ICC+ = 0.82). The use of a familiarization session before the assessment of the trunk extensor endurance may help the participant to learn, for example, the appropriate testing position needed to be adopted for the test and to better tolerate the fatigue feelings generated during the Biering-Sorensen test [10,56]. However, quite often it is impossible to carry out a familiarization session in practical (clinic, sports, fitness, and physical education) contexts. Although this circumstance affects the inter-session reliability estimate of the Biering-Sorensen test negatively, the magnitude of this effect is not large enough to consider this endurance measure inappropriate (according to the 0.8 cut-off score) for practical and research purposes.

Other qualitative moderators such as sex (males vs. females vs. males and females) and population of sample (children and adolescents vs. adults) also showed a tendency towards statistical significance. However, and similar to what was found for the moderator tool or equipment, these tendencies may be caused by the large imbalanced distribution shown in the number of cohorts included in each of the sub-categories of these moderators (e.g., males = 9 cohorts vs. females = 2 cohorts vs. males and females = 15 cohorts; adults = 23 cohorts vs. children and adolescents = 4 cohorts). Despite this, the sub-categories of these two moderators showed reliability estimates higher than 0.8, and hence the Biering-Sorensen test may be a suitable field-based test to be conducted in both sexes and youth and adult populations.

Finally, out of 103 studies that reported reliability coefficients, 75 studies induced reliability (72.8%). The number of studies that induced reliability according to our study is in concordance with the high rate of RI that usually characterizes research in general and which is around 70% of the studies conducted in the social and health sciences [73]. These findings reinforce the need to educate researchers that reliability is not an inherent characteristic of the instrument and that it should be analysed whenever a test is administered.

This RG meta-analysis presents some limitations that should be highlighted. First, the unadvised practice of inducing reliability reduced the number of studies included in this RG meta-analysis. Second, only studies in English or Spanish were included, which may have limited the number of potentially eligible studies. Third, the lack of important data reported by authors reduced the possibility of analysing their influence as potential moderating variables on the reliability coefficients. Such was the case of the age means and age standard deviation of the samples, as well as the sex distribution, the physical activity level, or the tester' profession and level of experience conducting the test. To finish, the scarce number of reliability estimates reported for the field tests other rather than the Biering-Sorensen test did not allow to analyse the moderator variables that may explain in part the variability found among the reliability coefficients for those tests.

3.5. Conclusions

The main findings of the current RG meta-analysis report that only the trunk extension endurance measure obtained through the Biering-Sorensen test may present sufficient scientific evidence in terms of inter- (ICC+ = 0.94, based on 8 cohorts) and intra-tester (intra-session (ICC+ = 0.88, based on 12 cohorts) and inter-session (ICC+ = 0.88, based on 27 cohorts)) reliability to justify its use (mainly in male adults) for practical and clinical purposes. Furthermore, this trunk extension endurance measure may be even more reliable when a familiarization session with the Biering-Sorensen testing procedure is previously carried out and a plumb-line is used to control the participant's position during the test.

None of the rest of the trunk extension endurance measures obtained through the other static (prone isometric chest raise test and prone double straight-leg raise test) and dynamic field-based (dynamic extensor endurance test) tests showed sufficient scientific evidence (in terms of reliability) to promote their use for practical and clinical purposes. Therefore, more reliability studies conducted in different cohorts are needed before the use of these field-based tests can be recommended to assess trunk extensor endurance in both practical and research settings.

Finally, the high RI rate found in this study (72.8%) suggests that researchers should be more aware of the fact that RI is an erroneous practice that should be eradicated because it can cause errors in the estimation of the measures used [74]. Thus, in future research, authors should report the reliability coefficient of the scores for the data being analysed, even when the focus of their research is not psychometric.

3.6. References

1. Abdelraouf, O.R.; Abdel-aziem, A.A. The relationship between core endurance and back dysfunction in collegiate male athletes with and without nonspecific low back pain. Int. J. Sports Phys. Ther. 2016, 11, 337–344.

2. Biering-Sorensen, F. Physical Measurements as Risk Indicators of Low-Back trouble over one-year period. Spine (Phila Pa 1976) 1984, 9, 106–119.

3. Luoto, S.; Heliövaara, M.; Hurri, H.; Alaranta, H. Static back endurance low-back pain and the risk of low-back pain. Clin. Biomech. 1995, 10, 323–324.

4. McGill, S.M.; Childs, A.; Liebenson, C. Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. Arch. Phys. Med. Rehabil. 1999, 80, 941–944.

5. Hibbs, A.E.; Thompson, K.G.; French, D.; Wrigley, A.; Spears, I. Optimizing performance by improving core stability and strength. Sport Med. 2008, 38, 995–1008.

6. McGill, S.M. Ultimate back Fitness and Performance, 4th ed.; Backfitpro Incorporated: Waterloo, ON, Canada, 2009.

7. García-Vaquero, M.P.; Barbado, D.; Juan-Recio, C.; López-Valenciano, A.; Vera-Garcia, F.J. Isokinetic trunk flexion-extension protocol to assess trunk muscle strength and endurance: Reliability, learning effect, and sex differences. J. Sport Health Sci. 2016, doi:10.1016/j.jshs.2016.08.011.

8. Gruther, W.; Wick, F.; Paul, B.; Leitner, C.; Posch, M.; Matzner, M.; Crevena, R.; Ebenbichler, G. Diagnostic accuracy and reliability of muscle strength and endurance measurements in patients with chronic low back pain. J. Rehabil. Med. 2009, 41, 613–619, doi:10.2340/16501977-0391.

9. Mayer, T.; Gatchel, R.; Betancur, J.; Bovasso, E. Trunk Muscle Endurance Measurement. Isometric Contrasted to Isokinetic Testing in Normal Subjects. Spine (Phila Pa 1976) 1995, 20, 920–927.

10. Juan-Recio, C.; López-Plaza, D.; Barbado, D.; García-Vaquero, M.P.; Vera-García, F.J. Reliability assessment and correlation analysis of 3 protocols to measure trunk muscle strength and endurance muscle strength and endurance. J. Sports Sci. 2018, 36, 357–364, doi:10.1080/02640414.2017.1307439.

11. Ito, T.; Shirado, O.; Suzuki, H.; Takahashi, M.; Kaneda, K.; Strax, T.E. Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. Arch. Phys. Med. Rehabil. 1996, 77, 75–79.

12. McIntosh, G.; Wilson, L.; Affleck, M.; Hall, H. Trunk and lower extremity muscle endurance: Normative data. J. Rehabil. Outcomes Meas. 1998, 2, 20–39.

13. Lanning, C.L.; Uhl, T.L.; Ingram, C.L.; Mattacola, C.G.; English, T.; Newsom, S. Baseline values of trunk endurance and hip strength in collegiate athletes. J. Athl. Train. 2006, 41, 427–434.

14. Hopkins, W.G. Measures of Reliability in Sports Medicine and Science. Sports Med. 2000, 30, 375–381, doi:10.2165/00007256-200030010-00001.

15. Riebe, D.; Ehrman, J.K.; Liguori, G.; Magal, M. ACSM's Guidelines for Exercsie Testing and Prescription, 10th ed.; Wolters Kluwer: Philadelphia, PA, USA, 2018.

16. Maier, T.; Gross, M.; Trösch, S.; Steiner, T.; Müller, B.; Bourban, P.; Schärer, C.; Hübner, K.; Wehrlin, J.; Tschopp, M. Manual Leistungsdiagnostik; BASPO: Magglingen, Switzerland, 2016.

17. Plowman, S.A.; Meredith, M.D. FITNESSGRAM®/ACTIVITYGRAM®. Reference Guide, 4th ed.; The Cooper Institute: Dallas, TX, USA, 2013.

18. Arab, M.A.; Salavati, M.; Ebrahimi, I.; Mousavi, M.E. Sensitivity, specificity and predictive value of the clinical trunk muscle endurance tests in low back pain. Clin. Rehabil. 2007, 21, 640–647, doi:10.1177/0269215507076353.

19. Moreau, C.E.; Green, B.N.; Johnson, C.D.; Moreau, S.R. Isometric back extension endurance tests: A review of the literature. J. Manip. Physiol. Ther. 2001, 24, 110–122 doi:10.1067/mmt.2001.112563.

20. Atkinson, G.; Nevill, A. Statistical Methods for Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. Sports Med. 1998, 26, 217–238, doi:10.2165/00007256-199826040-00002.

21. Kottner, J.; Audigé, L.; Brorson, S.; Donner, A.; Gajewski, B.J.; Hróbjartsson, A.; Roberts, C.; Shoukri, M.; Streiner, D.L. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J. Clin. Epidemiol. 2011, 64, 96–106, doi:10.1016/j.jclinepi.2010.03.002.

22. Hayen, A.; Dennis, R.J.; Finch, C.F. Determining the intra- and inter-observer reliability of screening tools used in sports injury research. J. Sci. Med. Sports 2007, 10, 201–210, doi:10.1016/j.jsams.2006.09.002.

23. Latimer, J.; Maher, C.G.; Refshauge, K.; Colaco, I. The reliability and validity of the Biering-Sorensen test in asymptomatic subjects and subjects reporting current or previous nonspecific low back pain. Spine (Phila Pa 1976) 1999, 24, 2085–2089, doi:10.1097/00007632-199910150-00004.

24. Larsson, H.; Tegern, M.; Monnier, A.; Skoglund, J.; Helander, C.; Persson, E.; Malm, C.; Broman, L.; Aasa, U. Content Validity Index and Intra- and Inter-Rater Reliability of a New Muscle Strength/Endurance Test Battery for Swedish Soldiers. PLoS ONE 2015, 10, e0132185, doi:10.1371/journal.pone.0132185.

25. Moffroid, M.; Haugh, L.; Haig, A.; Henry, S.; Pope, M. Endurance training of trunk extensor muscles. Phys. Ther. 1993, 73, 10–17.

26. Simmonds, M.J.; Olson, S.L.; Jones, S.; Hussein, T.; Lee, E.; Novy, D.; Radwan, H. Psychometric Characteristics and Clinical Usefulness of Physical Performance Tests in Patients with Low Back Pain. Spine (Phila Pa 1976) 1998, 23, 2412–

2421.

27. Hannibal, N.S., III; Plowman, S.A.; Looney, M.A.; Brandenburg, J. Reliability and Validity of Low Back Strength/Muscular Endurance Field Tests in Adolescents. J. Phys. Act. Health 2006, 3(Suppl. S2), 78–89, DOI:10.1123/jpah.3.s2.s78.

28. Balthazard, P.; de Goumoens, P.; Rivier, G.; Demeulenaere, P.; Ballabeni, P.; Deriaz, O. Manual therapy followed by specific active exercises versus a placebo followed by specific active exercises on the improvement of functional disability in patients with chronic non specific low back pain: A randomized controlled trial. BMC Musculoskelet. Disord. 2012, 13, 162, doi:10.1186/1471-2474-13-162.

29. Coorevits, P.; Danneels, L.; Cambier, D.; Ramon, H.; Vanderstraeten, G. Assessment of the validity of the Biering-Sorensen test for measuring back muscle fatigue based on EMG median frequency characteristics of back and hip muscles. J. Electromyogr. Kinesiol. 2008, 18, 997–1005, doi:10.1016/j.jelekin.2007.10.012.

30. Kankaanpää, M.; Laaksonen, D.; Taimela, S.; Kokko, S.M.; Airaksinen, O.; Hänninen, O. Age, sex, and body mass index as determinants of back and hip extensor fatigue in the isometric Sorensen back endurance test. Arch. Phys. Med. Rehabil. 1998, 79, 1069–1075, doi:10.1016/s0003-9993(98)90173-3.

31. Evans, K.; Refshauge, K.; Adams, R. Trunk muscle endurance tests: Reliability, and gender differences in athletes. J. Sci. Med. Sports 2007, 10, 447–455, doi:10.1016/j.jsams.2006.09.003.

32. Okada, T.; Huxel, K.C.; Nesser, T.W. Relationship between core stability, functional movement, and performance. J. Strength Cond. Res. 2011, 25, 252–261, doi:10.1519/JSC.0b013e3181b22b3e.

33. Vacha-Haase, T.; Kogan, L.R.; Thompson, B. Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. Educ. Psychol. Meas. 2000, 60, 509–522, doi:10.1177/00131640021970682.

34. Henson, R.K.; Thompson, B. Characterizing measurement error in scores across studies: Some recommendations for conducting ‘reliability generalization’

studies. Meas. Eval. Couns. Dev. 2002, 35, 113–126, doi:10.1080/07481756.2002.12069054.

35. Rodriguez, M.C.; Maeda, Y. Meta-analysis of coefficient alpha. Psychol. Methods 2006, 11, 306–322, doi:10.1037/1082-989X.11.3.306.

36. Sánchez-Meca, J.; López-López, J.A.; López-Pina, J.A. Some recommended statistical analytic practices when reliability generalization studies are conducted. Br. J. Math. Stat. Psychol. 2013, 66, 402–425, doi:10.1111/j.2044-8317.2012.02057.x.

37. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Ann. Intern. Med. 2009, 151, 264–269, doi:10.1371/journal.pmed.1000097.

38. Prinsen, C.A.C.; Mokkink, L.B.; Bouter, L.M.; Alonso, J.; Patrick, D.L.; de Vet, H.C.W.; Terwee, C.B. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual. Life Res. 2018, 27, 1147–1157, doi:10.1007/s11136-018-1798-3.

39. Mokkink, L.B.; de Vet, H.C.W.; Prinsen, C.A.C.; Patrick, D.L.; Alonso, J.; Bouter, L.M.; Terwee, C.B. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual. Life Res. 2018, 27, 1171–1179, doi:10.1007/s11136-017-1765-4.

40. Terwee, C.B.; Prinsen, C.A.C.; Chiarotto, A.; Westerman, M.; Patrick, D.; Alonso, J.; Bouter, L.M.; de Vet, H.C.W.; Mokkink, L.B. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. Qual. Life Res. 2018, 27, 1159–1170, doi:10.1007/s11136-018-1829-0.

41. Weir, J.P. Quantifying test-retest reliability using the intraclass correlation coefficient and the standard error of measurement. J. Strength Cond. Res. 2005, 19, 231–240, doi:10.1519/15184.1.

42. Salminen, J.J.; Maki, P.; Oksanen, A.; Pentti, J. Spinal mobility and trunk muscle strength in 15-year-old schoolchildren with and without low-back pain. Spine (Phila Pa 1976) 1992, 17, 405–411, doi:10.1097/00007632-199204000-00005.

43. Demoulin, C.; Sac, D.; Serre, L.; Maquet, D.; Crielaard, J.M.; Marc, V. Reproducibility and Suitability of Clinical Assessments of Trunk Flexor and

Extensor Muscles. J. Musculoskelet. Pain 2008, 16, 298–308, doi:10.1080/10582450802479651.

44. Sandrey, M.A.; Mitzel, J.G. Improvement in Dynamic Balance and Core Endurance after a 6-Week Core-Stability-Training Program in High School Track and Field Athletes. J. Sports Rehabil. 2013, 22, 264–271, doi:10.1123/jsr.22.4.264.

45. Terwee, C.B.C.; Mokkink, L.; Knol, D.; Ostelo, R.; Bouter, L.; de Vet, H. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. Qual. Life Res. 2012, 21, 651–657, doi:10.1007/s11136-011-9960-1.

46. Orwin, R.; Vevea, J. Evaluating coding decisions. In The Handbook of Research Synthesis and Meta-Analysis, 2nd ed.; Cooper, H., Hedges, L.V., Valentine, J.C., Eds.; Russel Sage Foundation: New York, NY, USA, 2009; pp. 177–203.

47. Borenstein, M.; Hedges, L.V.; Higgins, J.; Rothstein, H.R. Introduction to Meta-Analysis; John Wiley & Sons: Chichester, UK, 2009.

48. Sánchez-Meca, J.; Marín-Martínez, F. Confidence Intervals for the Overall Effect Size in Random-Effects Meta-Analysis. Psychol. Methods 2008, 13, 31–48, doi:10.1037/1082-989X.13.1.3.

49. Huedo-Medina, T.B.; Sánchez-Meca, J.; Marín-Martínez, F.; Botella, J. Assessing heterogeneity in meta-analysis: Q statistic or I 2 Index? Psychol. Methods 2006, 11, 193–206, doi:10.1037/1082-989X.11.2.193.

50. Lopez-Lopez, J.A.; Botella, J.; Sanchez-Meca, J.; Marin-Martinez, F. Alternatives for Mixed-Effects Meta-Regression Models in the Reliability Generalization Approach: A Simulation Study. J. Educ. Behav. Stat. 2013, 38, 443–469, doi:10.3102/1076998612466142.

51. López-López, J.A.; Marín-Martínez, F.; Sánchez-Meca, J.; Van den Noortgate, W.; Viechtbauer, W. Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. Br. J. Math. Stat. Psychol. 2014, 67, 30–48, doi:10.1111/bmsp.1200.

52. Borenstein, M.; Hedges, L.V.; Higgins, J.; Rothstein, H.R. Comprehensive Meta-Analysis 3.3.; Biostat, Inc: Englewood, NJ, USA, 2018.

53. Dedering, A.; Roos Af Hjelmsäter, M.; Elfving, B.; Harms-Ringdahl, K.; Németh,

G. Between-days reliability of subjective and objective assessments of back extensor muscle fatigue in subjects without lower-back pain. J. Electromyogr. Kinesiol. 2000, 10, 151–158, doi:10.1016/s1050-6411(00)00009-2.

54. Dedering, A.; Gnospelius, A.; Elfving, B. Reliability of Measurements of Endurance Time, Electromyographic Fatigue and Recovery, and Associations to Activity Limitations, in Patients with Lumbar Disc Herniation. Physiother. Res. Int. 2010, 15, 189–198, doi:10.1002/pri.457.

55. del Pozo-Cruz, B.; Mocholi, M.; del Pozo-Cruz, J.; Parraca, J.; Adsuar, J.; Gusi, N. Reliability and validity of lumbar and abdominal trunk muscle endurance tests in office workers with nonspecific subacute low back pain. J. Back Musculoskelet. Rehabil. 2014, 27, 399–408, doi:10.3233/BMR-140460.

56. Demoulin, C.; Boyer, M.; Duchateau, J.; Grosdent, S.; Jidovsteff, B.; Crielaard, J.M.; Vanderthommen, M. Is the Sørensen test valid to assess muscle fatigue of the trunk extensor muscles? J. Back Musculoskelet. Rehabil. 2016, 29, 31–40, doi:10.3233/BMR-150592.

57. Geldhof, E.; Cardon, G.; De Bourdeaudhuij, I.; Danneels, L.; Coorevits, P.; Vanderstraeten, G.; de Clercq, D. Effects of back posture education on elementary schoolchildren's back function. Eur. Spine J. 2007, 16, 829–839, doi:10.1007/s00586-006-0199-4.

58. Javadian, Y.; Behtash, H.; Akbari, M.; Taghipour-darzi, M.; Zekavat, H. The effects of stabilizing exercises on pain and disability of patients with lumbar segmental instability. J. Back Musculoskelet. Rehabil. 2012, 25, 149–155, doi:10.3233/BMR-2012-0321.

59. Jorgensen, K.; Nicolaisen, T. Two methods for determining trunk extensor endurance. A comparative study. Eur. J. Appl. Physiol. Occup. Physiol. 1986, 55, 639–644, doi:10.1007/bf00423210.

60. Juan-Recio, C.; Barbado, D.; López-Valenciano, A.; Vera-García, F.J. Test de campo para valorar la resistencia de los músculos del tronco. Apunt. Educ. Fis. Deport. 2014, 3, 59–68, doi:10.5672/apunts.2014-0983.es.(2014/3).117.06.

61. Keller, A.; Hellesnes, J.; Brox, J.I.I. Reliability of the isokinetic trunk extensor test, Biering-Sorensen test, and Astrand bicycle test: Assessment of intraclass

correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. Spine (Phila Pa 1976) 2001, 26, 771–777, doi:10.1097/00007632-200104010-00017.

62. Lin, K.-H.; Huang, Y.-M.; Tang, W.; Chang, Y.; Liu, Y.; Liu, C. Correlation of static and dynamic trunk muscle endurance and bat swing velocity in high school aged baseball players. Isokinet. Exerc. Sci. 2013, 21, 113–119, doi:10.3233/IES-130486.

63. Mannion, A.F.; Connolly, B.; Wood, K.; Dolan, P. The use of surface EMG power spectral analysis in the evaluation of back muscle function. J. Rehabil. Res. Dev. 1997, 34, 427–439.

64. Moffroid, M.; Reid, S.; Henry, S.; Haugh, L.; Ricamato, A. Some Endurance Measures in Persons With Chronic Low Back Pain. J. Orthop. Sports Phys. Ther. 1994, 20, 81–87, doi:10.2519/jospt.1994.20.2.81.

65. Ozcan Kahraman, B.; Sengul, Y.; Kahraman, T.; Kalemci, O. Developing a Reliable Core Stability Assessment Battery for Patients With Nonspecific Low Back Pain. Spine (Phila Pa 1976) 2016, 41, E844–E850, doi:10.1097/BRS.0000000000001403.

66. Palacin-Marin, F.; Esteban-Moreno, B.; Olea, N.; Herrera-Viedma, E.; Arroyo-Morales, M. Agreement between telerehabilitation and face-to-face clinical outcome assessments for low back pain in primary care. Spine (Phila Pa 1976) 2013, 38, 947–952, doi:10.1097/BRS.0b013e318281a36c.

67. de Souza, C.P.; Monteiro-Junior, R.S.; da Silva, E.B. Reliability of the endurance test for the erector spinae muscle. Fisioter. Mov. 2016, 29, 369–375, doi:10.1590/0103-5150.029.002.AO16.

68. Teyhen, D.S.; Shaffer, S.W.; Lorenson, C.L.; Wood, S.L.; Rogers, S.M.; Dugan, J.L.; Walker, M.J.; Childs, J.D. Reliability of lower quarter physical performance measures in healthy service members. US Army Med. Dep. J. 2011, 2, 37–49.

69. Waldhelm, A.; Li, L. Endurance tests are the most reliable core stability related measurements. J. Sports Health Sci. 2012, 1, 121–128, doi:10.1016/j.jshs.2012.07.007.

70. Denteneer, L.; Van Daele, U.; Truijen, S.; De Hertogh, W.; Meirte, J.; Stassijns, G.

Reliability of physical functioning tests in patients with low back pain: A systematic review. Spine J. 2018, 18, 190–207, doi:10.1016/j.spinee.2017.08.257.

71. Denteneer, L.; Stassijns, G.; De Hertogh, W.; Truijen, S.; Van Daele, U. Inter- and Intrarater Reliability of Clinical Tests Associated With Functional Lumbar Segmental Instability and Motor Control Impairment in Patients With Low Back Pain: A Systematic Review. Arch. Phys. Med. Rehabil. 2017, 98, 151–164.e6, doi:10.1016/j.apmr.2016.07.020.

72. Nunnally, J.C.; Bernstein, I.H. The Assessment of Reliability. Psychom. Theory 1994, 3, 248–292.

73. Sanchez-Meca, J.; Rubio-Aparicio, M.; Marin-Martinez, F.; Lopez-Pina, J.A. The phenomenon of reliability induction in social and health sciences. In Fifteenth Congress of Methodology of Behavioral Science; Palma de Mallorca, Spain, 2015.

74. Sánchez-Meca, J.; López-Pina, J.A.; López-López, J.A. Generalización de la fiabilidad: Un enfoque meta-analítico aplicado a la fiabilidad. Fisioterapia 2009, 31, 262–270, doi:10.1016/j.ft.2009.05.005.

ESTUDIO 4. Reliability of five trunk flexion and extension endurance field-based tests in high school-aged adolescents: ISQUIOS Program

3.1. Introduction

It has been suggested that deficits in trunk extensor and flexor endurance and imbalances between trunk muscle groups may have short and long-term negative consequences in low back health [1,2] and athlete movement competency [3,4]. This circumstance has led to the field-based assessment of trunk muscle endurance becoming common practice during childhood and adolescence, mainly in educational (physical education (PE) classes) and sport settings.

Some field-based tests have been described to assess trunk extensor and flexor muscle endurance, which may be grouped into two main categories:

a) Isometric trunk extension (e.g.: Biering-Sorensen (BS) test [5], Prone Isometric Chest Raise (PICR) test [6] and Prone Double Straight-leg Raise (PDSR) test [7]) and flexion (e.g.: Ito test [6], Flexor Endurance (FE) test [8], Isometric Trunk Flexion endurance (ITF) test [9], Plank Isometric Hold (PIH) test [10] and Side Bridge right (SB-R) and left (SB-L) test [8]) endurance tests, which involve maintaining a position against gravity for as long as possible.
b) Dynamic trunk extension (e.g.: Dynamic Extensor Endurance (DEE) test [11]) and flexion (e.g.: Bench Trunk Curl-Up (BTC) test [12], Partial Curl-Up (PCU) test [13], Curl-Up (CU) test [14]) endurance tests, which consist of performing as many repetitions as possible in a given time or with a certain cadence until exhaustion.

These field-based tests have been considered operationally valid by medical (ACSM), sport (Swiss Olympic Medical Centre) and educational (Cooper Institute) organizations to assess trunk muscles endurance based on anatomical knowledge and findings presented in electromyographic [15–17] and biomechanical [18–20] studies. It should be highlighted that a single (i.e. all-out) field-based test able to concurrently quantify trunk muscles isometric and dynamic endurance has not been described in the literature (to the authors' knowledge). Therefore, a comprehensive assessment of trunk endurance capability is required for each muscle group (e.g. flexors and extensors), by selecting at least one isometric and dynamic test [18,21].

Reliability is a technical property of a measure or test that provides information regarding the consistency and reproducibility of given values in repeated trials [22]. The degree of reliability in a measure may be affected by the inter- and intra-individual variability in its scores within the sample object of study [23]. Consequently, reliability is a population dependent property (e.g.: children and adolescents, adults, athletes). Large inter-individual differences and fluctuations over short-time periods in strength and endurance scores have been documented in youth [24] and attributed, among other factors, to periods of rapid changes in growth and maturation and fluctuations in their mood state (e.g.: inter-day differences in the psychological readiness to perform a maximal effort to exhaustion) [22,25]. Therefore, before promoting the use of these trunk endurance field-based tests in youth, the reliability of their measures must be confirmed in this population [22]. A recently published meta-analysis of reliability of the measures obtained from trunk extension endurance field-based tests concluded that, in terms of inter-session reliability, there is no compelling evidence that supports their use in children and adolescents[26]. In particular, only 3 studies were identified that provided inter-session reliability scores for the measures obtained from the BS and DEE in children [9,27] and adolescents [28] with all of them reporting intra-class correlation coefficient (ICC) scores higher than 0.80. Concerning the reliability of measures obtained from trunk flexion endurance field-based tests, the evidence available is also very limited in youth. Only 1 study (to the authors' knowledge) has explored the inter-tester reliability of the dynamic endurance measure obtained from the CU test in children (10-12 years old) [29], whereas 5 studies have determined the inter-session reliability of the measures from PIH [30] ITF [9], isometric PCU [28] and BTC [25] tests in children and adolescents, showing ICC scores higher than 0.75.

Another limitation of the literature is that most of the studies that have explored the reliability of the trunk endurance field-based tests [9,27–30], although not all [25], have exclusively used the ICC as a criterion of reliability. The use of the ICC as the sole statistical outcome of reliability, apart from being affected by sample heterogeneity, only provides information regarding how well the observed value retains the true rank order of subjects but does not allow the quantification of either the extent of the measurement error, the presence of systematic bias or the minimum change needed for a specific outcome to consider that an improvement or decrease after an intervention program may be real or true (out of the random error threshold). Therefore, contemporary statistical approaches in which the most powerful statistical methods, such as the typical (random) percentage error and the minimal detectable change at a 95% confidence interval (MDC_{95}) were included could be more useful in practical settings and for clinical goals.

Therefore, the purpose of the present study was to explore the inter-session reliability of the measures obtained from 2 popular trunk extension (BS and DEE tests) and 3 trunk flexion (Ito, SB [R and L] and BTC tests) endurance field-based tests using a contemporary statistical approach in high school-aged adolescents of 4 different age-based grade levels. The null hypothesis is that unlike dynamic field-based tests, isometric trunk extensor and flexor muscle endurance measures would show acceptable (for clinical purposes and goal setting) and stronger reliability than dynamic measures, independent of the participants' age-based grade, due primarily to their easy assessment procedures [25,31,32].

3.2. Materials and methods

3.2.1. Participants

A total of 241 adolescents were initially invited from 3 different high schools of the Region of Murcia (Spain) to participate in this study (convenience sample). The exclusion criteria were: a) known medical problems or episodes of low back pain over the last 3 months (reported by the PE teachers), b) not having provided the required signed written informed consent (by both the parents/guardians and students) before the start of the study, c) missing 1 testing session during the data collection phase or d) involvement in structured strength exercise programs during the time of the study. Participants were asked not to perform strenuous exercises in the 24 h before each assessment session.

A comprehensive verbal description of the nature and purpose of the study and the experimental risks was given to the students and their parents/guardians and PE teachers. The study was conducted according to the Declaration of Frontera and the protocol was fully approved by the Review Committee for Research Involving Human Subjects at the University of Murcia (Spain) (ID: 1920/2018).

Finally, a sample of 208 (age 14.4 ± 1.2 years (mean ± SD), range 12-18 years, 52.4% girls) high school students (age-based grades (n): 1st grade (53), 2nd grade (43), 3rd grade (65) and 4th grade (47)) completed this study (Table 25). Thirty-three students were removed from the initial sample of 241 adolescents based on the exclusion criteria (10 students (3 boys and 7 girls) reported a history of low back pain, 14 students (6 boys and 8 girls) did not provide the required signed informed consent before the start of the study and 9 students (3 boys and 6 girls) did not attend one or both testing sessions).

Table 25. Anthropometric characteristics of the participants (mean ± SD) (n=208).

	1st grade		2nd grade		3rd grade		4th grade	
	M	F	M	F	M	F	M	F
Sample	24	29	22	21	37	28	12	35
Age (years)	13.1±0.6	13±0.8	13.9±0.7	13.7±0.4	14.8±0.5	14.9±0.5	15.9±0.3	16.1±0.5
Body mass (kg)	50.1±9.7	56.2±12	61.8±14.8	54.3±9.9	61.7±13.9	56.8±10.4	70.5±15.4	62.3±11.3
Stature (cm)	157.1±6.8	156.3±6.6	167.3±8.2	159.7±6.5	168.7±7.4	160.1±5.7	176±7.6	161.1±5.5
BMI (kg/m^2)	20.2±3.1	22.9±4.5	21.9±4.4	21.3±4.1	21.5±4.2	22.1±3.9	22.6±4.2	23.9±3.7

M = males, F = females, kg = kilogram, cm = centimetre, BMI = body mass index.

3.2.2. Study design and procedure

A test-retest design was used to determine the inter-session reliability (both grouped (all participants pooled in the same data set) and separately by age-grade) of the trunk endurance measures obtained from 5 field-based tests during PE classes. Two field-based tests were selected to assess isometric (BS test) [5] and dynamic (DEE test) [11] trunk extensor endurance (Figure 17 A and B) whereas 3 field-based tests were selected to assess isometric (Ito and SB-R and SB-L tests) [6,8] and dynamic (BTC test) [12] trunk flexor endurance (Figure 17 C, D and E). Taking into account that there is no single field-based test available to simultaneously quantify the isometric and dynamic endurance of the trunk muscles, it was necessary to use at least one isometric and dynamic test for each muscle group [21].

Since only 2 sessions of 60 min per age-based grade were provided by PE teachers from each high school, a time-efficient testing procedure was designed and 5 researchers were enrolled to enable the assessment of participants on 2 different occasions with a 7-days rest interval between them. The same protocol was consistently followed in all the testing sessions conducted in the 3 high schools that took part in this study. For each age grade, both testing sessions were administered at the same time of the day during PE classes and under the direct supervision of the same 5 researchers, who were sports science specialists with more than 5 years of experience in neuromuscular performance assessments. Each researcher was responsible during both testing sessions for the same field-based test. Furthermore, the participants' testing sequence and environmental factors (e.g.: temperature, setting) were the same during the 2 testing sessions (Temperature: 22˚C, relative humidity: 50-60%).

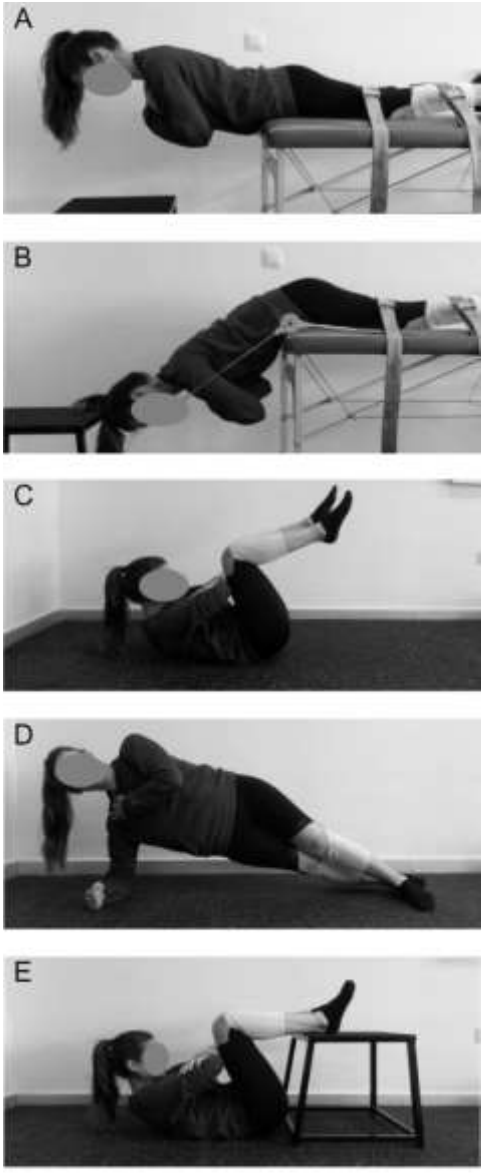


Figure 17. Trunk endurance field-based tests. A and B: trunk extension endurance field-based tests (BS test and DEE test). C, D, and E: trunk flexion endurance trunk field-based tests (Ito test, SB test, and BTC test). BS = Biering-Sorensen test, DEE = Dynamic Extensor Endurance test, SB = Side Bridge test, BTC = Bench Trunk Curl-Up test.

At the start of the 2 testing sessions, all the participants received comprehensive instructions for the tests, and their questions regarding the protocols were answered. In each testing session, all participants completed first their usual warm-up, which was led by their PE teachers and consisted of 6-10 min of low-to moderate-intensity (self-perceived) running (including forward/backward movements and side-stepping) and general mobilization (i.e., arm circles, leg kicks) followed by 4-6 min of static stretching. Afterwards, 10 minutes was provided to all participants to freely familiarize them with the test procedures, all participants were instructed to perform a minimum of 2 sets of 3 to 5 repetitions of each dynamic trunk endurance test and 3 sets of 5 s for the isometric endurance test. Once the participants understood and freely practised the tests, participants randomly performed the 5 field-based trunk endurance tests with a 5-min rest between each test. Due to the above-mentioned time constraints, a randomized circuit approach was used to carry out all the tests (both in the test and the retest session). Five different stations were set (one for each trunk endurance test). In groups of 4-6 participants were randomise assigned to one of the 5 stations. At each station, participants alternatively performed the test so that while one of them was carrying out the test the others were resting. After 8 min, groups were moved to their next station (clockwise) until all of them were completed.

An extendable goniometer (Lafayette Instrument Co, Lafayette, IN, USA) was used to ensure the correct joint position was maintained during the tests. During the performance of all field-based tests, participants were strongly encouraged verbally to maintain the position as long as possible or to perform the maximum number of repetitions as possible. Participants did not receive any feedback on performance until the end of the study.

3.2.3. Trunk extensor endurance field-based tests

Biering-Sorensen test

The isometric endurance of the trunk extensor musculature was assessed through the BS test [5]. The test started with the participant in a prone position with the lower body resting on a test bench and the anterior superior iliac spine aligned at the edge of the test bench. The lower body was attached to the test bench by 2 inextensible straps (knees and ankles). In the starting position, the upper body rested with both forearms placed on a chair. During the test, the upper body was maintained in a horizontal position (0 degrees of hip flexion) with arms crossed on the chest while holding the head in a neutral position (Figure 17A). The test consisted of maintaining the trunk in the described position for as long as possible, until exhaustion, or until participants lost the correct position more than 3 times. A loss of the correct position during the execution of the test was identified when participants flexed their hips more than 10º (determined using extendable goniometer). The test duration was recorded in seconds.

Dynamic Extensor Endurance test

The dynamic endurance of the trunk extensor muscles was assessed through the DEE test [11]. Participants were located in the same position as the BS test. In the starting position, hip flexion of 45º was performed and both forearms rested on a chair. During the test, participants had to extend the trunk horizontally and then return to the initial position with arms crossed on the chest (Figure 17B). Participants were asked to carry out the maximum repetitions possible in 60 seconds. Only those repetitions that were performed correctly were counted, that is, those in which the trunk was fully extending (horizontally), and in which the head touched the chair when flexing the hip. The hip flexion during the test was controlled through a static reference (extendable goniometer).

3.2.4. Trunk flexor endurance field-based tests

Ito test

The isometric endurance of the trunk flexor muscles was assessed through the Ito test [6]. Participants were placed in a supine position with hips and knees flexed 90° (extendable goniometer) and arms interlaced with hands grasping the opposite elbow. From this position, participants performed a trunk flexion ("curl-up") until they touched their thighs with their elbows, the scapulae did not touch the mat and the head was in a neutral position. The test consisted of maintaining this position for as long as possible, until exhaustion. The test ended when the scapulae came in contact with the mat, recording the test duration in seconds.

The original test was modified to normalize the range of motion to the participants' characteristics and thus avoid hip and lower back flexion ("sit-up") [19,20]. For this, before starting the test, the participant was placed in the aforementioned initial position, and then the subject performed a trunk flexion until the scapulae did not touch the mat. From this position, the tester approached the participant's legs towards their elbows, until they came into contact (Figure 17C). Then, the tester held the legs in this new position while the participant rested before starting the test. This leg position was maintained throughout the test.

Side Bridge test

The isometric endurance of the trunk lateral flexor musculature was assessed through the SB-R and SB-L test [8]. Participants were placed in a lateral position on their side (supported by either the dominant and non-dominant arm depending on the side tested) with legs extended. The participants were supported on their elbow and feet, the top foot was placed ahead of the lower foot (with 90º elbow flexion and the arm perpendicular to the mat) while bridging their hips off the mat to maintain an aligned body position. The uninvolved arm was held across the chest with the hand placed on the opposite shoulder (Figure 17D). The test finished when the subject lost the aligned postural position, and the duration recorded in seconds. Both sides were tested with the dominant side always examined first.

Bench Trunk Curl-Up test

The dynamic endurance of the trunk flexor muscles was assessed through the BTC test [12]. Participants were placed in a supine position with hips and knees flexed at 90° (extendible goniometer) and resting on a bench. The arms were crossed with the hands grasping the opposite elbow (Figure 17E). From this position, participants performed a trunk flexion ("curl-up") until they touched their thighs with their elbows, the scapulae did not touch the mat and the head was in a neutral position and then returned to the initial position. Just like the Ito test, a modification of the original test was performed to avoid hip and lower back flexion ("sit-up"), approaching the participant's legs towards their elbows, until they came into contact. The test consisted of performing the maximum number of repetitions possible in 2 minutes. Only those repetitions that were performed correctly were counted, that is, those in which the elbows touched the thighs in the flexing of the trunk, and in which the head touched the mat when lowering the trunk.

3.2.5. Statistical Analyses

Data are presented as mean ± SD. The distribution of each endurance measure was examined with the Shapiro-Wilk normality test and all measures were shown to be normally distributed.

In line with the current consensus regarding the determination of reliability in human performance-based studies, the following three aspects were assessed for each test using grouped and grade-specific measures [33,34]: 1) relative reliability, 2) presence (or not) of systematic bias between testing sessions and 3) precision of measurements (absolute reliability). Furthermore, the sensitivity of each test was also assessed [35]. The relative reliability was examined by the intra-class correlation coefficient ($ICC_{3,1}$) and an ICC >0.70 was considered acceptable [36].

The assessment of systematic bias between testing sessions was carried out via the Bayesian paired t-test (with a Cauchy distribution with spread r set to 0.707). The BF_{10} was interpreted using the evidence categories suggested by Lee and Wagenmakers [37]: <1/100 = extreme evidence for H_0, from 1/100 to 1/30 = very strong evidence for H_0, from 1/30 to 1/10 = strong evidence for H_0, from 1/10 to 1/3 = moderate evidence for H_0, from 1/3 to 1 anecdotal evidence for H_0, from 1 to 3 = anecdotal evidence for H_1, from 3 to 10 = moderate evidence for H_1, from 10 to 30 = strong evidence for H_1, from 30 to 100 = very strong evidence for H_1, >100 extreme evidence for H_1. The median and the 95% central credible interval (CI) of the posterior distribution of the standardized effect size (δ) (i.e. the population version of Cohen's d) were also calculated for each of the paired-comparisons carried out. Magnitudes of the posterior distribution of the standardized effect size were classified as: trivial (<0.2), small (0.2 – 0.6), moderate (0.6 – 1.2), large (1.2 – 2.0) and very large (2.0 – 4.0) [38]. Only those pairwise comparisons that showed at least strong evidence for supporting the alternative hypothesis (BF_{10} >10), an error percentage <10 (which indicates great stability of the numerical algorithm that was used to obtain the result) and δ >0.6 (at least moderate) were considered robust to describe significant differences.

A Bland-Altman plot was built for each trunk endurance measure to graphically show mean bias and 95% limits of agreement. Heteroscedasticity was assessed using a Bayesian correlation coefficient (Pearson´s rho) between the means of the participant's test and retest scores and the absolute differences between the participant's test and retest scores [22]. To qualitatively interpret the size of the Bayesian correlation coefficients, the thresholds defined by Hinkle, Wiersma & Jurs [39] for Behavioural Sciences were followed: from 0 to 0.3 = negligible correlation, from 0.3 to 0.5 = low correlation, from 0.5 to 0.7 = moderate, from 0.7 to 0.9 = high, from 0.9 to 1 = very high. Only correlations higher than 0.5 (at least moderate) were considered relevant for these sub-analyses.

The precision of measurement was determined using the typical percentage error and the minimal detectable change at a 95% confidence interval (MDC_{95}) using the Hopkins' spreadsheet [40]. The typical percentage error (coefficient of variation (CV_{TE})) was calculated using the log-transformed data via the following formula: $100\,(e^S - 1)$, where s is the typical error of measurement (TEM) (SD of the difference between testing session 1 and testing session 2 divided by $\sqrt{2}$). Logarithmic transformations of the data were performed and used to reduce the possible heteroscedasticity of the raw data [41]. To interpret the CV_{TE} values, the current study used the arbitrary value suggested by Weir and Vincent [42] and Hopkins [22] with an analytical goal of 10% or below to consider a test as demonstrating good inter-session reliability. The MDC_{95} was calculated as follow: $CV_{TE} \times 1.96 \times \sqrt{2}$.

The sensitivity of each test was assessed while comparing the smallest worthwhile percentage change (SWC) with the CV_{TE}. The SWC was determined by multiplying the pure between-testing sessions SD by 0.2 ($SWC_{0.2}$), which corresponds to a small effect, 0.6 ($SWC_{0.6}$), which corresponds to a moderate effect and 1.2 ($SWC_{1.2}$), which corresponds to a large effect. If the CV_{TE} was lower than the SWC, the test was rated as "good"; if the CV_{TE} was similar to the SWC, the rating was "OK"; and if the CV_{TE} was higher than the SWC, a rating of "marginal" was given [35].

Statistical analysis was performed using the JASP computer software Version 0.11.1 (JASP Team, Amsterdam, The Netherlands) and the online Hopkins' spreadsheet (www.sportsci.org).

3.3. Results

The grouped and age grade-specific descriptive statistics and reliability values for the trunk flexion and extension endurance measures obtained from the 5 field-based tests selected are shown in tables 26 and 27, respectively.

Table 26. Descriptive statistics (mean ± SD) and reliability scores (mean and 90% confidence interval) of the trunk extension endurance measures obtained from the 2 field-based tests selected.

	Grouped data (n = 208)	Grade specific data			
		1er grade (n = 53)	2nd grade (n = 43)	3rd grade (n = 65)	4th grade (n = 47)
BS test (s)					
- Testing session 1	132.5±54.6	139.8±58.9	133.4±55.6	135.3±56.9	122.9±47.3
- Testing session 2	130.5±52.0	137.9±61.1	129.9±56.4	133.2±51	122.1±45
- Systematic bias (%)	-1.3 (-4.0 to 1.6)	-2.8 (-9.2 to 4)	-3.9 (-10.9 to 3.7)	-0.6 (-5.4 to 4.5)	0.3 (-4.8 to 5.7)
- $ICC_{3,1}$	0.86 (0.82 to 0.89)	0.89 (0.81 to 0.94)	0.82 (0.70 to 0.89)	0.84 (0.76 to 0.90)	0.88 (0.80 to 0.93)
- CV_{TE} (%)	16 (14.6 to 17.9)	16.7 (13.6 to 21.8)	20.3 (16.7 to 26.2)	15.8 (13.5 to 19.4)	14.1 (11.7 to 17.9)
- MDC_{95} (%)	44.5 (40.4 to 49.6)	46.3 (37.7 to 60.4)	56.3 (46.3 to 72.6)	43.8 (37.4 to 53.7)	39.1 (32.4 to 49.6)
DEE test (rep)					
- Testing session 1	39.0±10.6	36.4±8.2	41.8±9.7	43.6±11.5	33.2±8.9
- Testing session 2	43.3±12.6	37.8±9.9	47.4±13.2	49.1±12.7	37.9±12
- Systematic bias (%)	10 (7.1 to 12.9)*	2.6 (-3.9 to 9.5)	12.3 (7.2 to 17.6)*	12.2 (6.9 to 17.7)*	13.1 (6.4 to 20.1)*
- $ICC_{3,1}$	0.77 (0.72 to 0.82)	0.58 (0.37 to 0.74)	0.81 (0.70 to 0.89)	0.73 (0.59 to 0.82)	0.77 (0.64 to 0.86)
- CV_{TE} (%)	15.3 (14 to 17)	17.9 (14.8 to 22.8)	12.8 (10.7 to 16.1)	15.2 (12.9 to 18.5)	17.4 (14.5 to 21.9)
- MDC_{95} (%)	42.4 (38.7 to 47.2)	49.6 (41 to 63.2)	35.5 (29.6 to 44.6)	42.1 (35.7 to 51.3)	48.2 (40.2 to 60.7)

Table 26. Descriptive statistics (mean ± SD) and reliability scores (mean and 90% confidence interval) of the trunk extension endurance measures obtained from the 2 field-based tests selected.

	Grouped data (n = 208)	Grade specific data			
		1er grade (n = 53)	2nd grade (n = 43)	3rd grade (n = 65)	4th grade (n = 47)

*: there was at least a strong evidence (BF_{10} >10) to support the alternative hypothesis (H_1: the presence of relevant inter-session differences) with at least a moderate effect size (δ >0.6). ICC = intraclass correlation coefficient, CV_{TE} = typical percentage error, MDC = minimum detectable change, BS = Biering-Sorensen test, DEE = Dynamic Extensor Endurance test

The grouped and age grade-specific relative reliability scores (i.e. $ICC_{3,1}$) found in this study for all the trunk endurance measures were higher than 0.7 (except the ICC scores for the DEE test in participants in 1st grade) and hence, they may be considered as acceptable according to the thresholds previously reported in the literature [36].

Table 27. Descriptive statistics (mean ± SD) and reliability scores (mean and 90% confidence interval) of the trunk flexion endurance measures obtained from the 2 field-based tests selected.

	Grouped data (n = 208)	Grade specific data			
		1er grade (n = 53)	2nd grade (n = 43)	3rd grade (n = 65)	4th grade (n = 47)
Ito test					
- Testing session 1	160.1±142.5	128.5±69.1	135.8±125.7	199.2±183.1	150.2±122.3
- Testing session 2	165.4±147.4	123.6±76.8	149.3±140.6	204.7±184.9	154.8±125.3
- Systematic bias (%)	1.3 (-3.1 to 5.9)	-9.3 (-17.5 to -0.2)	3 (-7.1 to 14.2)	3.7 (-4.8 to 12.8)	2.6 (-5 to 10.8)
- $ICC_{3,1}$	0.94 (0.92 to 0.95)	0.92 (0.83 to 0.97)	0.95 (0.90 to 0.97)	0.95 (0.91 to 0.97)	0.93 (0.88 to 0.95)
- CV_{TE} (%)	24.4 (21.9 to 27.5)	17.2 (13.2 to 25.2)	25 (20 to 33.6)	26.7 (22.3 to 33.4)	24.4 (20.4 to 30.3)
- MDC_{95} (%)	67.5 (60.8 to 76.1)	47.6 (36.6 to 70.7)	69.3 (55.4 to 93.1)	74 (61.8 to 92.6)	67.6 (56.5 to 83.9)
SB-R test					
- Testing session 1	57.3±29.4	51.3±29.8	62.4±31.4	58±24.9	59.1±33.9
- Testing session 2	56.4±27.4	50.5±28.1	60±31.1	59.2±23.7	58.5±30.8
- Systematic bias (%)	1.7 (-1.7 to 5.3)	4.6 (-5.8 to 16)	-1.9 (-10 to 6.9)	4.2 (-0.1 to 8.7)	2.9 (-3.5 to 9.8)
- $ICC_{3,1}$	0.90 (0.87 to 0.92)	0.85 (0.75 to 0.91)	0.91 (0.84 to 0.95)	0.93 (0.90 to 0.96)	0.92 (0.87 to 0.95)
- CV_{TE} (%)	19.7 (17.9 to 22)	29.4 (24 to 38.1)	20.8 (16.7 to 27.6)	13.3 (11.3 to 16.2)	18.3 (15.3 to 23.1)
- MDC_{95} (%)	54.6 (49.5 to 60.9)	81.5 (66.52 to 105.6)	57.6 (46.3 to 76.5)	36.8 (31.3 to 44.9)	50.7 (42.4 to 64)

Table 27. Descriptive statistics (mean ± SD) and reliability scores (mean and 90% confidence interval) of the trunk flexion endurance measures obtained from the 2 field-based tests selected.

	Grouped data (n = 208)	Grade specific data			
		1er grade (n = 53)	2nd grade (n = 43)	3rd grade (n = 65)	4th grade (n = 47)
SB-L test					
- Testing session 1	57.3±29.7	52.3±30.2	64.3±34.5	58.9±25.9	54.3±29.3
- Testing session 2	55.5±27.4	51.2±28.6	61.8±31.2	58.2±24	50.9±26
- Systematic bias (%)	-1.2 (-4.3 to 2.1)	-0.3 (-7 to 6.9)	-2.3 (-10.8 to 7)	0.3 (-4.7 to 5.5)	-2.2 (-8.6 to 4.8)
- $ICC_{3,1}$	0.92 (0.90 to 0.94)	0.93 (0.88 to 0.96)	0.92 (0.86 to 0.96)	0.91 (0.86 to 0.94)	0.93 (0.88 to 0.96)
- CV_{TE} (%)	18.5 (16.8 to 20.7)	19.1 (15.8 to 24.4)	21.6 (17.4 to 29)	16.2 (13.7 to 19.8)	18.8 (15.5 to 23.9)
- MDC_{95} (%)	51.4 (46.6 to 57.4)	52.9 (43.8 to 67.6)	59.8 (48.2 to 80.4)	44.9 (37.9 to 54.8)	52.1 (42.9 to 66.2)
BTC test (rep)					
- Testing session 1	57.5±18.8	54.6±18.7	63.8±24.3	54.6±16.9	61.7±16
- Testing session 2	64.9±19.8	60.4±19.6	70.6±24.9	65.9±19.3	66.7±15.8
- Systematic bias (%)	13.6 (11.3 to 15.9)*	11.3 (5.9 - 16.9)*	11.9 (8.4 to 15.6)*	21.7 (17.6 to 25.9)*	8.5 (4.4 to 12.8)*
- $ICC_{3,1}$	0.90 (0.87 to 0.92)	0.90 (0.83 - 0.95)	0.94 (0.90 to 0.96)	0.91 (0.86 to 0.95)	0.89 (0.82 to 0.94)
- CV_{TE} (%)	11.3 (10.3 to 12.6)	12.4 (10.2 - 15.9)	10.3 (8.4 to 12.4)	10.4 (8.8 to 12.7)	10.2 (8.5 to 12.9)
- MDC_{95} (%)	31.4 (28.5 to 34.9)	34.4 (28.3 to 44.1)	29.4 (23.3 to 34.4)	28.8 (2.4 to 35.2)	28.3 (23.5 to 35.7)

*: there was at least a strong evidence (BF_{10} >10) to support the alternative hypothesis (H_1: the presence of relevant inter-session differences) with at least a moderate effect size (δ >0.6). ICC = intraclass correlation coefficient, CV_{TE} = typical percentage error, MDC = minimum detectable change, SB-L = Side Bridge Left test, SB-R = Side Bridge Right test, BTC = Bench Trunk Curl-Up test.

The separate Bayesian paired t-test analyses carried out to explore potential inter-session differences (systematic bias) only revealed the existence of at least a strong evidence in favour of the alternative hypothesis (H1) with at least a moderate effect size (δ >0.6) in the trunk extension and flexion endurance measures obtained from their respective dynamic field-based tests (DEE and BTC tests). These relevant inter-session differences for the DEE and BTC tests were found across all the age-based grades. Bland-Altman plots illustrate the differences between the two testing sessions (y-axis) and the mean value of each of the paired measurements (x-axis) for each trunk endurance field-based test (Figure 18 (grouped data). Dashed lines illustrate the systematic bias and random error forming the 95% limits of agreement.

The heteroscedasticity coefficients for the grouped and age grade-specific measures obtained through the 5 trunk endurance field-based tests were not relevant ($r < 0.5$) (except the heterogeneity correlation scores found for the Ito tests ($r = 0.58$) in the 2nd grade) (Figure 18).

For the grouped and age grade-specific endurance measures obtained from the field-based tests, the CV_{TE} scores ranged from 12.8% (DEE test) to 20.3% (BS test) and from 10.2% (BTC test) to 29.4% (SB-R test) for the trunk extensor and flexor muscles, respectively. These CV_{TE} scores may be considered as poor according to the 10% cut-off score used in this study. For the MDC_{95}, their scores ranged from 35.5% (DEE test) to 56.3% (BS test) and from 28.3% (BTC test) to 81.5% (SB-R test) for the trunk flexion and extension endurance measures, respectively.

As displayed in Table 28, all field-based tests were sensitive to detect moderate and large changes (independent of the participants' age-based grade), but none of the tests was considered sensitive enough to detect small changes.

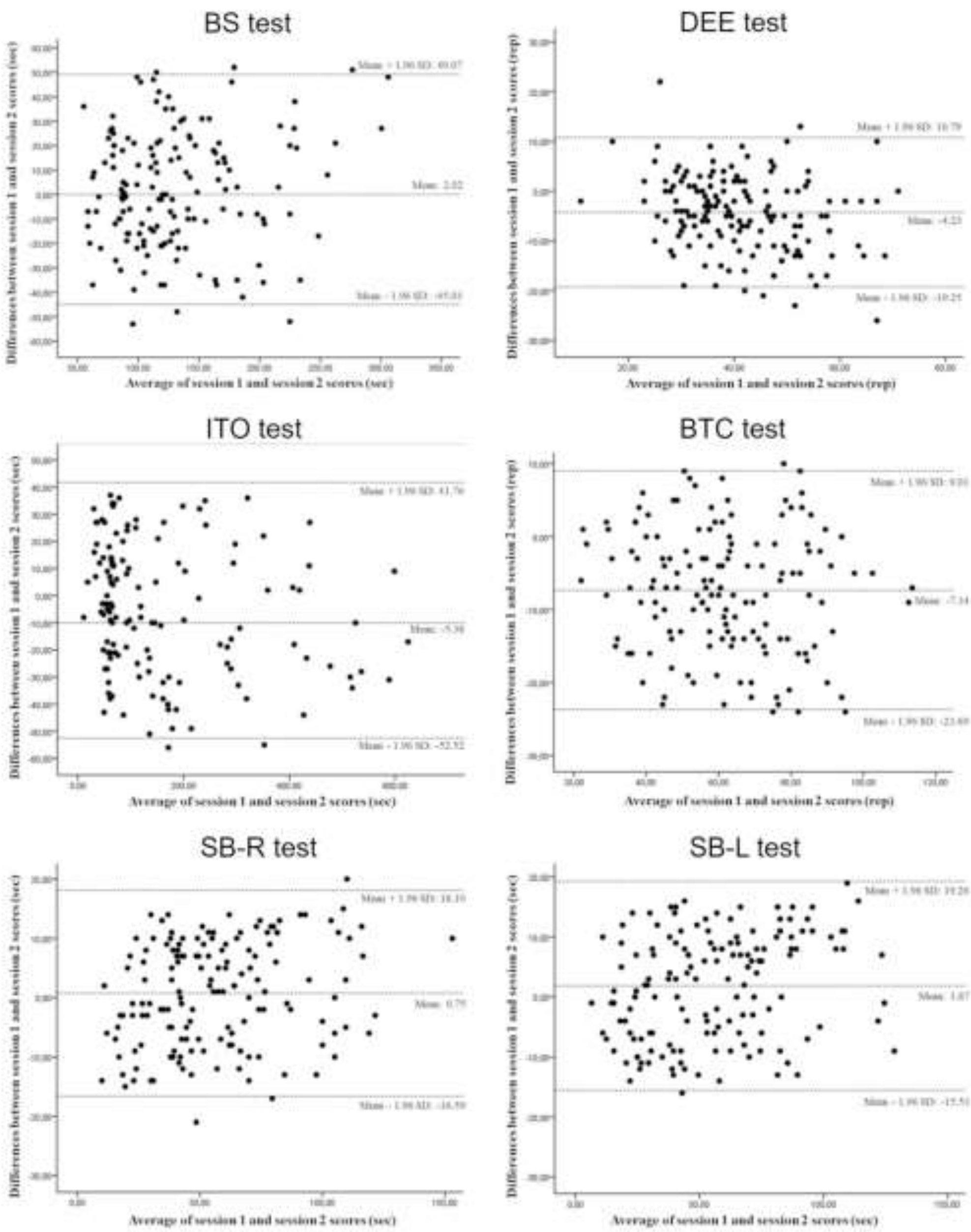


Figure 18. Bland-Altman plots for the trunk endurance field-based tests (grouped data (n = 208). Trunk extension endurance field-based tests (BS test and DEE test). Trunk flexion endurance trunk field-based tests (Ito test, SB test, and BTC test). BS = Biering-Sorensen test, DEE = Dynamic Extensor Endurance test, SB = Side Bridge test, BTC = Bench Trunk Curl-Up test.

Table 28. The sensitivity of the field-based trunk endurance tests to detect small, moderate and large changes. Rating of sensitivity is also provided.

	Grouped data (n = 208)	Grade specific data			
		1er grade (n = 53)	2nd grade (n = 43)	3rd grade (n = 65)	4th grade (n = 47)
Trunk extension endurance field-based tests					
BS test					
- $SWC_{0.2}$	7.5 (marginal)	8.9 (marginal)	8 (marginal)	6.9 (marginal)	7.2 (marginal)
- $SWC_{0.6}$	15 (OK)	26.7 (good)	24 (good)	20.7 (good)	21.6 (good)
- $SWC_{1.2}$	45 (good)	53.4 (good)	48 (good)	41.4 (good)	43.2 (good)
DEE test					
- $SWC_{0.2}$	5.4 (marginal)	3.9 (marginal)	5.01 (marginal)	4.6 (marginal)	5.9 (marginal)
- $SWC_{0.6}$	10.8 (good)	11.7 (marginal)	15.03 (marginal)	13.8 (marginal)	17.7 (OK)
- $SWC_{1.2}$	32.5 (good)	23.4 (good)	30.1 (good)	27.6 (good)	35.4 (good)
Trunk flexion endurance field-based tests					
Ito test					
- $SWC_{0.2}$	18.1 (marginal)	10.8 (marginal)	19.3 (marginal)	21.4 (marginal)	16.3 (marginal)
- $SWC_{0.6}$	36.2 (good)	32.4 (good)	57.9 (good)	64.2 (good)	48.9 (good)
- $SWC_{1.2}$	108.6 (good)	64.8 (good)	115.8 (good)	128.4 (good)	97.8 (good)
SB-R test					
- $SWC_{0.2}$	11.5 (marginal)	12.8 (marginal)	12.4 (marginal)	9.7 (marginal)	11.7 (marginal)
- $SWC_{0.6}$	23 (good)	38.4 (good)	37.2 (good)	29.1 (good)	35.1 (good)
- $SWC_{1.2}$	69 (good)	76.8 (good)	74.4 (good)	58.2 (good)	70.2 (good)
SB-L test					
- $SWC_{0.2}$	12.3 (marginal)	13 (marginal)	14.1 (marginal)	9.7 (marginal)	13.1 (marginal)
- $SWC_{0.6}$	24.6 (good)	39 (good)	42.3 (good)	29.1 (good)	39.3 (good)
- $SWC_{1.2}$	73.8 (good)	78 (good)	84.6 (good)	58.2 (good)	78.6 (good)
BTC test					
- $SWC_{0.2}$	6.7 (marginal)	7.1 (marginal)	1.1 (marginal)	6.5 (marginal)	5.6 (marginal)
- $SWC_{0.6}$	13.4 (good)	21.4 (good)	3.2 (marginal)	19.5 (good)	16.8 (good)
- $SWC_{1.2}$	40.2 (good)	42.9 (good)	6.5 (marginal)	39 (good)	33.6 (good)

SWC = smallest worthwhile percentage change, BS = Biering-Sorensen test, DEE = Dynamic Extensor Endurance test, SB-L = Side Bridge Left test, SB-R = Side Bridge Right test, BTC = Bench Trunk Curl-Up test.

3.4. Discussion

The main purpose of this study was to examine the inter-session reliability of the trunk flexor and extensor endurance measures obtained from 5 common field-based tests during PE classes in a large sample of high school-aged adolescents. The analyses of the inter-session reliability and sensitivity carried out with the grouped and age grade-specific trunk flexor and extensor endurance measures consistently reported similar results across the 4 age-based grades. Therefore, both the relative and absolute reliability and also the sensitivity scores obtained from each field-based test using the measures from the whole sample of participants (i.e. grouped data set) may be considered as robust (based on the large sample size (n = 208)) and generalizable criteria to be used when assessing and monitoring trunk endurance in high school-aged adolescents (i.e. independent of the age-grade). Thus, the findings of the present study indicate that all trunk flexor and extensor endurance measures demonstrate acceptable (ICC > 0.7) relative reliability. However, significant inter-session differences (i.e. systematic bias) were found in the measures obtained from the DEE and BTC tests. Likewise, the precision of the measurement of each field-based test was poor (CV_{TE} <10%) with the MDC_{95} revealing that changes higher than 42% for trunk extension endurance tests and 31.4% for trunk flexion endurance tests after an intervention are required to indicate a significant change above measurement error. Finally, the sensitivity analyses conducted revealed that all tests were sensitive enough to detect moderate to large changes in trunk muscle endurance.

Concerning the relative reliability, similar results (ICC > 0.75) were found in previous studies for the measures of BS [28], DEE [27,28] and BTC [25]. The relative reliability scores of the measures obtained from the Ito and SB tests cannot be compared with the finding reported in previous studies because to the best of the authors' knowledge, this is the first study that has explored inter-session reliability in adolescents. The acceptable relative reliability results found in the present study for the measures of the 5 field-based tests selected might have been positively impacted by the large inter-individual variability observed in the trunk endurance scores (i.e. the 5 tests reported SDs larger than 20%). The heterogeneity documented for the participants' trunk endurance scores may be partially attributed to the large inter-individual difference (regarding level [magnitude of change], tempo [rate of change] and timing [onset of change]) in maturity status that is often found in adolescents within a given age group (up to 15 cm and 21 kg in the stature and body mass, respectively) [43,44].

The analysis of the presence of systematic bias between testing sessions revealed that this phenomenon only appeared in the trunk endurance measures obtained from the DEE and BTC tests, independent of the age-based grade of the participants. These results were similar to the findings of Moya-Ramón et al. [25] who also found statistically significant inter-session differences in the endurance measure obtained through the BTC test in adolescents. The significant inter-session differences in BTC and DEE measures have been also confirmed by the 95% limits of agreement. For example, when an adolescent performed 60 repetitions on the BTC, on the retest he could perform as high as 60 + 9 =

69 repetitions or as low as 60 - 23.7 = 36.3 repetitions. An explanation for these significant inter-session differences may be based on the dynamic nature of both the DEE and BTC tests. For example, the execution of these two tests requires participants to perform as many repetitions as possible with a certain cadence until exhaustion, which would entail the need of completing a pre-assessment familiarization session to learn their testing procedures when individuals who will be tested have little or no experience with them [25,31,45]. Given the time constraints in this study all participants were only allowed to freely practice the tests for ten minutes in each testing session. The main reason behind the implementation of this short and unsupervised familiarization protocol in each testing session was that all the field-based tests selected seem to present simple testing procedures. Thus, a short and unsupervised familiarization protocol was thought to be enough to avoid significant inter-session differences in the trunk endurance scores. However, this short and unsupervised familiarization protocol may have not been sufficient for participants to learn how to maintain the specific cadence of movements required during the execution of the DEE and BTC tests [25,31,42,46]. Consequently, in the first session, some participants may have performed more repetitions that did not meet the criteria for a successful repetition than during the second test, due to the relatively short familiarization with the testing procedures of the dynamic tests at that time [21,26,47]. The fact that all participants had limited or no experience with the trunk endurance field-based tests selected may explain the significant inter-session differences found for the measures from DEE and BTC and these differences might be attributed to a possible learning effect, which was consistent throughout the 4 age-based grades. Therefore, these findings suggest that in adolescents (12-18 years) with limited or no experience (independent of their age) with the DEE and BTC tests, that the use of a longer and/or supervised familiarization protocol or the inclusion of an additional testing session is highly recommended to minimize the learning effects that were observed in the analyses of the inter-session differences of their measures. By the end of familiarization, practitioners should be confident that each adolescent has fully understood the testing procedures and has acquired a satisfactory proficiency to correctly perform both tests (focusing the attention in the ability to maintain the specific cadence of movements required during the execution of these tests). Another factor that may have contributed to the systematic bias found in the endurance measures from the dynamic field-based tests is higher motivation in the second session by some participants (information based on testers' comments). Although the trunk endurance scores (time and repetitions) achieved

by participants in the five tests selected were not revealed by the testers until the end of the data collection phase, it might have been possible that some participants had mentally counted the number of repetitions completed during the two dynamic tests in the first session. Thus, and after having compared their results among each other, in the second session, some participants exhibited significant motivation to achieve better scores in the dynamic tests than those obtained by their peers during the first session [9,22,25,31,48]. This circumstance was not as evident during the execution of the isometric tests because their final scores were determined in seconds rather than the number of repetitions.

The results of the present study also demonstrate that the precision of measurement of each field-based test could be categorized as poor. In particular, the CV_{TE} of each field-based test (CV_{TE} ranged from 11.3 to 24.4) exceeds the widely accepted 10% cut-off value to consider the magnitude of the measurement error as satisfactory, for both clinical and practice goals in healthy populations [22,42]. Only the CV_{TE} of the endurance measure from the BTC (CV_{TE} = 11.3) test approached this 10% cut-off score. Moya-Ramón et al. [25] also found poor CV_{TE} scores for the BTC test in boys (17.2%) aged 14-18. The precision of measurement of the rest of the field-based tests cannot be directly compared with previous studies as this is the first time that CVs have been calculated in an adolescent population. A plausible reason that may explain the poor CV_{TE} scores of the trunk endurance measures found in this study may be also attributed to the fact that these physically demanding field-based tests are substantively influenced by motivation as no external/internal indicator of maximal effort is available. Consequently, these findings recommend that PE teachers convey to adolescents before carrying out the tests the vital importance that a maximal effort is required to achieve a true or real estimation of the trunk endurance capability. During the execution of each test, strong verbal encouragement is required for the adolescents to maintain the position required in the isometric tests for as long as possible, and to perform as many repetitions as possible in the dynamic tests. In turn, this may be also contributing to an improvement in the reliability scores of the trunk endurance measures recorded.

The MDC_{95} value might be regarded as the minimum amount of change that needs to be observed, at either the group or individual level, for it to be considered a real or true change with a 95% level of certainty [49,50] (i.e. greater than the random measurement error). Thus, and according to the results of this study, changes higher than 31.4, 42.4, 44.5, 51.4, 54.6 and 67.5% for the BTC, DEE, BS, SB-L, SB-R and Ito tests (respectively) after an intervention may be considered as real or true with a 95% level of certainty [51].

The magnitude of the changes needed to consider any training intervention as effective to improve trunk muscle endurance of the BTC test (31%) for high-school-aged adolescents may be acceptable as previous studies have reported improvements of approximately 40% in trunk endurance after having completed a 6-week training program [52,53]. However, for the DEE, BS, SB-L, SB-R, and Ito tests these reported "improvements" of 40% in trunk endurance may not be true but simply measurement error. These findings are in line with the results of the sensitivity analyses that showed that these filed-based tests exhibited a good ability to detect moderate ($SWC_{0.6}$) to large ($SWC_{1.2}$) changes in their measures.

While the results of this study have provided information regarding the inter-session reliability of trunk muscle endurance measures obtained from five common field-based tests in an applied environment, limitations to the study must be acknowledged. Only high school-aged students were selected in this study and hence, the generalizability of the results to other populations (e.g.: adults) cannot be ascertained. Similarly, whether the trunk endurance measures would be as reliable in a population of injured adolescents (e.g.: low back pain) must be considered, although these field-based tests are generally performed in healthy, uninjured populations. Finally, for the dynamic field-based tests, the use of a longer and/or supervised familiarization protocol or the inclusion of an additional testing session might have minimized the learning effects that were observed in the analysis of the inter-session differences. Future applied studies should explore the inter-session reliability of the trunk endurance measures obtained through filed-based tests using supervised familiarization sessions in young populations. Besides, the lack of evidence regarding the inter-tester reliability of the measures obtained from the trunk extension endurance field-based tests in this cohort warrants further investigation.

3.5. Conclusion

To conclude, the findings from this study indicate that the trunk endurance measures obtained from five popular field-based tests (BTC, DEE, BS, SB-L, SB-R, and Ito) present poor inter-session reliability scores in high school-aged students (12-16 years). Even though all trunk endurance measures exhibited good relative reliability scores (ICC > 0.75), the precision of measurement of each field-based test was poor (CV_{TE} > 10%). Furthermore, inter-session systematic bias was observed in the dynamic trunk endurance field-based tests (BTC and DEE). The MDC_{95} results indicate that changes higher than 31.4, 42.4, 44.5, 51.4, 54.6, and 67.5% for the BTC, DEE, BS, SB-L, SB-R, and Ito tests after an intervention may be considered as real or true with a 95% level of certainty. Only the BTC test obtained a MDC_{95} low enough (31.4%) to be considered as acceptable for high school-aged students as previous studies have reported improvements of approximately 40% in trunk endurance after having completed a 6-week training program. The use of supervised familiarization sessions before performing the tests (focus the attention on dynamic tests) and strongly encourage adolescents to do always their maximum effort in each test may help improve the reliability scores shown in this study.

3.6. References

1. Steffens, D.; Maher, C.G.; Pereira, L.S.M.; Stevens, M.L.; Oliveira, V.C.; Chapple, M.; et al. Prevention of low back pain. A systematic review and meta-analysis. JAMA Intern. Med. 2016, 176(2), 199–208, doi: 10.1001/jamainternmed.2015.7431.

2. Gomes-Neto, M.; Lopes, J.M.; Conceição, C.S.; Araujo, A.; Brasileiro, A.; Sousa, C.; et al. Stabilization exercise compared to general exercises or manual therapy for the management of low back pain: A systematic review and meta-analysis. Phys. Ther. Sport. 2017, 23, 136–42, doi: 10.1016/j.ptsp.2016.08.004.

3. De Blaiser, C.; Roosen, P.; Willems, T.; Danneels, L.; Bossche, L.V.; De Ridder, R. Is core stability a risk factor for lower extremity injuries in an athletic population? A systematic review. Phys Ther Sport. 2018, 30, 48–56, doi: 10.1016/j.ptsp.2017.08.076.

4. Huxel Bliven, K.C.; Anderson, B.E. Core Stability Training for Injury Prevention. Sports Health. 2013, 5(6), 514–522, doi: 10.1177/1941738113481200.

5. Biering-Sorensen, F. Physical Measurements as Risk Indicators of Low-Back trouble over one-year period. Spine (Phila Pa 1976). 1984, 9(2), 106–119, doi: 10.1097/00007632-198403000-00002.

6. Ito, T.; Shirado, O.; Suzuki, H.; Takahashi, M.; Kaneda, K.; Strax, T.E. Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. Arch. Phys. Med. Rehabil. 1996, 77, 75–79, doi: 10.1016/S0003-9993(96)90224-5.

7. McIntosh, G.; Wilson, L.; Affleck, M.; Hall, H. Trunk and lower extremity muscle endurance: Normative data. J. Rehabil. Outcomes Meas. 1998, 2(4), 20–39.

8. McGill, S.M.; Childs, A.; Liebenson, C. Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. Arch. Phys. Med. Rehabil. 1999, 80(8), 941–944, doi: 10.1016/S0003-9993(99)90087-4.

9. Geldhof, E.; Cardon, G.; De Bourdeaudhuij, I.; Danneels, L.; Coorevits, P.; Vanderstraeten, G.; et al. Effects of back posture education on elementary schoolchildren's back function. Eur. Spine J. 2007, 16(6), 829–839, doi:

10.1007/s00586-006-0199-4.

10. Bliss, L.S.; Teeple, P. Core stability: the centerpiece of any training program. Curr. Sports Med. Rep. 2005, 4(3), 179–183, doi: 10.1097/01.CSMR.0000306203.26444.4e.

11. Luoto, S.; Heliövaara, M.; Hurri, H.; Alaranta, H. Static back endurance low-back pain and the risk of low-back pain. Clin. Biomech. 1995, 10(6), 323–324, doi: 10.1016/0268-0033(95)00002-3.

12. Knudson, D.; Johnston, D. Validity and Reliability of a Bench Trunk-Curl Test of Abdominal Endurance. J. Strength Cond. Res. 1995, 9(3), 165–169.

13. Jetté, M.; Sidney, K.; Cicutti, N. A critical analysis of sit-ups: a case for the partial curl-up as a test of abdominal muscular endurance. CAHPER J. 1984, 51(1), 4–9.

14. Robertson, L.D.; Magnusdottir, H. Evaluation of criteria associated with abdominal fitness testing. Res. Q. Exerc. Sport. 1987, 58(4), 355–9, doi: 10.1080/02701367.1987.10608112.

15. Coorevits, P.; Danneels, L.; Cambier, D.; Ramon, H.; Vanderstraeten, G. Assessment of the validity of the Biering-Sorensen test for measuring back muscle fatigue based on EMG median frequency characteristics of back and hip muscles. J. Electromyogr Kinesiol. 2008, 18(6), 997–1005, doi: 10.1016/j.jelekin.2007.10.012.

16. Müller, R.; Strässle, K.; Wirth, B. Isometric back muscle endurance: An EMG study on the criterion validity of the Ito test. J. Electromyogr. Kinesiol. 2010, 20(5), 845–850, doi: 10.1016/j.jelekin.2010.04.004.

17. Monfort-Pañego, M.; Vera-García, F.J.; Sánchez-Zuriaga, D.; Sarti-Martínez, M.Á. Electromyographic Studies in Abdominal Exercises: A Literature Synthesis. J. Manipulative Physiol. Ther. 2009, 32(3), 232–244, doi: 10.1016/j.jmpt.2009.02.007.

18. Moreau, C.E.; Green, B.N.; Johnson, C.D.; Moreau, S.R. Isometric back extension endurance tests: A review of the literature. J. Manipulative Physiol. Ther. 2001, 24(2), 110–122, doi: 10.1067/mmt.2001.112563.

19. Juan-Recio, C.; Lopez-Vivancos, A.; Moya, M.; Sarabia, J.M.; Vera-García, F.J. Short-term effect of crunch exercise frequency on abdominal muscle endurance. J. Sports Med. Phys. Fitness. 2015, 55(4), 280–289.

20. Juan-Recio, C.; Barbado, D.; López-Valenciano, A.; Vera-García, F.J. Test de campo para valorar la resistencia de los músculos del tronco. Apunt Educ. Fis. y Deport. 2014, 3(117), 59–68, doi: 10.5672/apunts.2014-0983.es.(2014/3).117.06.

21. Juan-Recio, C.; López-Plaza, D.; Barbado, D.; García-Vaquero, M.P.; Vera-García, F.J.

Reliability assessment and correlation analysis of 3 protocols to measure trunk muscle strength and endurance muscle strength and endurance. J. Sports Sci. 2018, 36(4), 357–364, doi: 10.1080/02640414.2017.1307439.

22. Hopkins, W.G. Measures of Reliability in Sports Medicine and Science. Sport Med. 2000, 30(5), 375–381, doi: 10.2165/00007256-200030050-00006.

23. Matheson, G.J. We need to talk about reliability: Making better use of test-retest studies for study design and interpretation. PeerJ. 2019, 7(e6918), doi: 10.7717/peerj.6918.

24. Lloyd, R.S.; Oliver, J.L.; Faigenbaum, A.D.; Myer, G.D.; De Ste Croix, M.B. Chronological age vs. biological maturation: implications for exercise programming in youth. J Strength Cond Res. 2014, 28(5), 1454–1464, doi: 10.1519/JSC.0000000000000391.

25. Moya-Ramón, M.; Juan-Recio, C.; Lopez-Plaza, D.; Vera-Garcia, F.J. Dynamic trunk muscle endurance profile in adolescents aged 14-18: Normative values for age and gender differences. J. Back Musculoskelet. Rehabil. 2018, 31(1), 155–162, doi: 10.3233/BMR-169760.

26. Martínez-Romero, M.T.; Ayala, F.; De Ste Croix, M.; Vera-Garcia, F.J.; Sainz de Baranda, P.; Santonja-Medina, F.; et al. A Meta-Analysis of the Reliability of Four Field-Based Trunk Extension Endurance Tests. Int J Environ Res Public Health. 2020, 17(9), 3088, doi: 10.3390/ijerph17093088.

27. Hannibal III, N.S.; Plowman, S.A.; Looney, M.A.; Brandenburg, J. Reliability and Validity of Low Back Strength/Muscular Endurance Field Tests in Adolescents. J Phys Act Health. 2006, 3(Suppl. 2), 78–89, doi: 10.1123/jpah.3.s2.s78.

28. Lin, K-H.; Huang, Y-M.; Tang, W.; Chang, Y.; Liu, Y.; Liu, C. Correlation of static and dynamic trunk muscle endurance and bat swing velocity in high school aged baseball players. Isokinet. Exerc. Sci. 2013, 21(2), 113–119, doi: 10.3233/IES-130486 10.3233/IES-130486.

29. Patterson, P.; Bennington, J.; de la Rosa, T. Psychometric properties of child- and teacher-reported curl-up scores in children ages 10-12 years. Res. Q. Exerc. Sport. 2001, 72(2), 117–124, doi: 10.1080/02701367.2001.10608941.

30. Boyer, C.; Tremblay, M.; Saunders, T.; McFarlane, A.; Borghese, M.; Lloyd, M.; et al. Feasibility, Validity, and Reliability of the Plank Isometric Hold as a Field-Based Assessment of Torso Muscular Endurance for Children 8-12 Years of Age. Pediatr. Exerc. Sci. 2013, 25(3), 407–422, doi: 10.1123/pes.25.3.407.

31. Brotons-Gil, E.; García-Vaquero, M.; Peco-González, N.; Vera-Garcia, F. Flexion-

Rotation Trunk Test To Assess Abdominal Muscle Endurance: Reliability, Learning Effect, and Sex Differences. J. Strength Cond. Res. 2013, 27(6), 1602–1608, doi: 10.1519/JSC.0b013e31827124d9.

32. De Blaiser, C.; De Ridder, R.; Willems, T.; Danneels, L.; Roosen. P. Reliability and validity of trunk flexor and trunk extensor strength measurements using handheld dynamometry in a healthy athletic population. Phys. Ther. Sport. 2018, 34, 180–186, doi: 10.1016/j.ptsp.2018.10.005.

33. Hopkins, W.G.; Marshall, S.W.; Batterham, A.M.; Hanin, J. Progressive Statistics for Studies in Sports Medicine and Exercise Science. Med. Sci. Sports Exerc. 2009, 41(1), 3–12, doi: 10.1249/MSS.0b013e31818cb278.

34. Kottner, J.; Audigé, L.; Brorson, S.; Donner, A.; Gajewski, B.J.; Hróbjartsson, A.; et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J. Clin. Epidemiol. 2011, 64(1), 96–106, doi: 10.1016/j.jclinepi.2010.03.002.

35. Impellizzeri, F.M.; Marcora, S.M. Test validation in sport physiology: Lessons learned from clinimetrics. Int. J. Sports Physiol. Perform. 2009, 4(2), 269–277, doi: 10.1123/ijspp.4.2.269.

36. Baumgartner, T.A.; Chung, H. Confidence Limits for Intraclass Reliability Coefficients. Meas. Phys. Educ. Exerc. Sci. 2001, 5(3), 179–188, doi: 10.1207/S15327841MPEE0503_4.

37. Lee, M.D.; Wagenmakers, E.J. Bayesian cognitive modeling: A practical course. Cambirdge University Press: United Kingdom, 2013.

38. Batterham, A.M.; Hopkins, W.G. Making meaningful inferences about magnitudes. Int. J. Sports Physiol. Perform. 2006, 1, 50–57.

39. Hinkle, D.; Wiersma, W.; Jurs, S. Applied Statistics for the Behavioral Sciences. 5th ed. Houghton Mifflin: Boston, USA, 2003.

40. Hopkins, W.G. Spreadsheets for Analysis of Validity and Reliability. Sportscience. 2015, 19, 36–42. Available from: sportsci.org/2015/ValidRely.htm

41. Atkinson, G.; Nevill, A. Statistical Methods for Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. Sport Med. 1998, 26(4), 217–238, doi:10.2165/00007256-199826040-00002.

42. Weir, J.P.; Vincent, W.J. Statistics in Kinesiology. 5th ed. Human Kinetics Publ Inc: Champaign, USA, 2020.

43. Evans, K.; Refshauge, K.; Adams, R. Trunk muscle endurance tests: Reliability, and

gender differences in athletes. J. Sci. Med. Sport. 2007, 10(6), 447–455, doi: 10.1016/j.jsams.2006.09.003.

44. Radnor, J.M.; Oliver, J.L.; Waugh, C.M.; Myer, G.D.; Lloyd, R.S. The influence of maturity status on muscle architecture in school-aged boys. Pediatr. Exerc. Sci. 2020, 32(2), 89–96, doi: 10.1123/pes.2019-0201.

45. De Blaiser, C.; De Ridder, R.; Willems, T,; Danneels, L.; Bossche, L.V.; Palmans, T.; et al. Evaluating abdominal core muscle fatigue: Assessment of the validity and reliability of the prone bridging test. Scand. J. Med. Sci. Sport. 2018, 28(2), 391–399, doi: 10.1111/sms.12919.

46. Strand, S.L.; Hjelm, J.; Shoepe, T.C.; Fajardo, M.A. Norms for an Isometric Muscle Endurance Test. J. Hum. Kinet. 2014, 40(1), 93–102, doi: 10.2478/hukin-2014-0011.

47. Demoulin, C.; Boyer, M.; Duchateau, J.; Grosdent, S.; Jidovsteff, B.; Crielaard, J-M.; et al. Is the Sørensen test valid to assess muscle fatigue of the trunk extensor muscles? J. Back Musculoskelet. Rehabil. 2016, 29, 31–40, doi: 10.3233/BMR-150592.

48. Narayan, A.; Steele-Johnson, D. Relationships between prior experience of training, gender, goal orientation and training attitudes. Int. J. Train. Dev. 2007, 11(3), 166–180, doi: 10.1111/j.1468-2419.2007.00279.x.

49. Beninato, M.; Portney, L.G. Applying concepts of responsiveness to patient management in neurologic physical therapy. J. Neurol. Phys. Ther. 2011, 35(2), 75–81, doi: 10.1097/NPT.0b013e318219308c.

50. Portney, L.G. Foundations of Clinical Research: Applications to Evidence-Based Practice. 4th ed. FA Davis Company: Philadelphia, USA, 2020.

51. Furlan, L.; Sterr, A. The applicability of standard error of measurement and minimal detectable change to motor learning research - A behavioral study. Front. Hum. Neurosci. 2018, 12(95), doi: 10.3389/fnhum.2018.00095.

52. González-Gálvez, N.; Marcos-Pardo, P.J.; Carrasco-Poyatos, M. Functional improvements after a pilates program in adolescents with a history of back pain: A randomised controlled trial. Complement. Ther. Clin. Pract. 2019, 35, 1–7, doi: 10.1016/j.ctcp.2019.01.006.

53. Granacher, U.; Schellbach, J.; Klein, K.; Prieske, O.; Baeyens, J.; Muehlbauer, T. Effects of core strength training using stable versus unstable surfaces on physical fitness in adolescents: a randomized controlled trial. BMC Sport Sci. Med. Rehabil. 2014, 6(40), 1–11, doi: 10.1186/2052-1847-6-40.

ESTUDIO 5. Effect of maturation on trunk muscle endurance in 8 to 16-year-old children and adolescents from the Region of Murcia (Spain): ISQUIOS Program

3.1. Introduction

The presence of back pain (BP) in children and adolescents is increasingly common [1]. Although most report absence of pain, an increase in the prevalence of BP is observed with age in children and adolescents [2]. BP prevalence within the last 12 months in children and adolescents was reported to be 62% in children between 10 to 13 years old, increasing to 83% in young people aged 17 to 19 years [3]. The main consequences of the presence of BP during childhood and adolescence include school absenteeism, the need for medical attention and medication [4], along with difficulty in performing different activities of daily life (i.e. standing in a queue, carrying a backpack or doing physical activities) [1].

A recent study concluded that pubertal development and linear growth were associated with spinal pain. It was observed that boys and girls with more advanced pubertal development and those undergoing greater growth experienced increased spinal pain frequency and duration [5]. Hence, it appears that puberty is the time of a rapid increase in the prevalence values of BP, being this prevalence higher among girls due to their early onset of puberty (12 years of age in females and around 14 years in males) [6–8].

Among the changes that occur during the growth spurt phase, it is well known that differential growth rates exist between the legs and the trunk, with the long bones of the legs experiencing peak growth ahead of the shorter bones of the trunk [9,10]. However, these changes are not always followed by a similar onset and rate of muscle development involved in the spine stability (e.g., multifidus, transversus abdominis, external and internal abdominal obliques, erector spinae, quadratus lumborum, and rectus abdominis). Therefore, the phase of peak height velocity (PHV) could be considered a particularly vulnerable period compared with the episodes before or after that stage due to rapid mechanical loading changes on the spine [5,6,11,12]. This momentary situation could alter the trunk muscles endurance [13,14], the alignment of the spinal curvatures [15], and leading to the appearance of BP [16], among others.

In fact, trunk muscles endurance is considered a risk factor associated with BP [16,17]. Several studies have concluded that reduced endurance of trunk extensor muscles, as well as trunk flexor muscles, was associated with low back pain (LBP) in adolescents [18–20]. These links between poor trunk muscles endurance and BP justify their assessment during childhood and adolescences by researchers, physiotherapists, physical education (PE) teachers and strength and conditioning coach to monitor their development during growth stages, determine a treatment/training effectiveness or develop prevention strategies for back disorders [21–23].

Maturational sex-related differences before the adolescent spurt do not differ substantially in body height, body mass, girth, muscle strength, bone width, or skinfold thickness [8,24]. However, with the onset of puberty, female and male endocrinology differs; females secreting more estrogen and males secreting more testosterone. During the early part of the puberty, girls are temporarily taller and heavier because of their earlier growth spurt [8,10]. Girls soon lose the size advantage as the growth spurt of boys, and boys achieve greater stature and muscle mass [8,10,24]. Due to these physical changes associated with puberty, girls during the PHV tend to perform worse in tasks that involve weight-bearing or endurance. In contrast, boys generally tend to have a superior athletic performance during PHV, particularly in activities that require strength, speed, or power [8,25,26].

Considering that a significant inter-individual variation exists for the level (magnitude of change), timing (onset of change), and tempo (rate of change) of maturation [27], the use of CA should be limited when assessing physical characteristics in children and adolescents. Therefore, this study aimed to analyse and compare the influence of CA and maturational stage on trunk muscles endurance in males and females school-aged. Understanding how trunk endurance measures evolve from pre-PHV to post-PHV, may not only help sports science specialists to distinguish between changes caused by growth and maturation and those mediated by physical training, but also to design tailored maturational stage-based training programs to prevent BP and optimize trunk motor performance. The studies carried out to date in children and adolescents [13,14,28–31] have evaluated the endurance of the trunk muscles differentiating by CA, so this study could give a more realistic vision of the changes that occur during maturation given the existing inter-individual differences among schoolchildren of the same CA.

3.2. Method

3.2.1. Design

It was a cross-sectional study. All measurements of trunk muscles endurance of students were collected before participate in a postural and physical fitness program called “ISQUIOS Program”. The study was conducted during the first term of the 2017–2018 and 2018-2019 school years.

3.2.2. Participants

A total of 994 students (primary education, n=548; secondary education, n=446) were initially invited from 10 and 4 different schools and high schools, respectively, of the Region of Murcia (Spain) to participate in this study (convenience sample). The exclusion criteria were: a) to have a diagnosed spine pathology or important physical injury which limited the correct performance of the tests, b) not to return signed the informed consent (both from parents/guardians and students) before the start of the study, c) to miss the testing session during the data collection, and d) to participate in structured conditioning exercise programs during the time of the study.

A comprehensive verbal description of the nature and purpose of the study and the experimental risks was given to the students and their parents/guardians and PE teachers. The protocol was fully approved by the Review Committee for Research Involving Human Subjects at the University of Murcia (Spain) (ID: 1920/2018) and according to the Declaration of Helsinki.

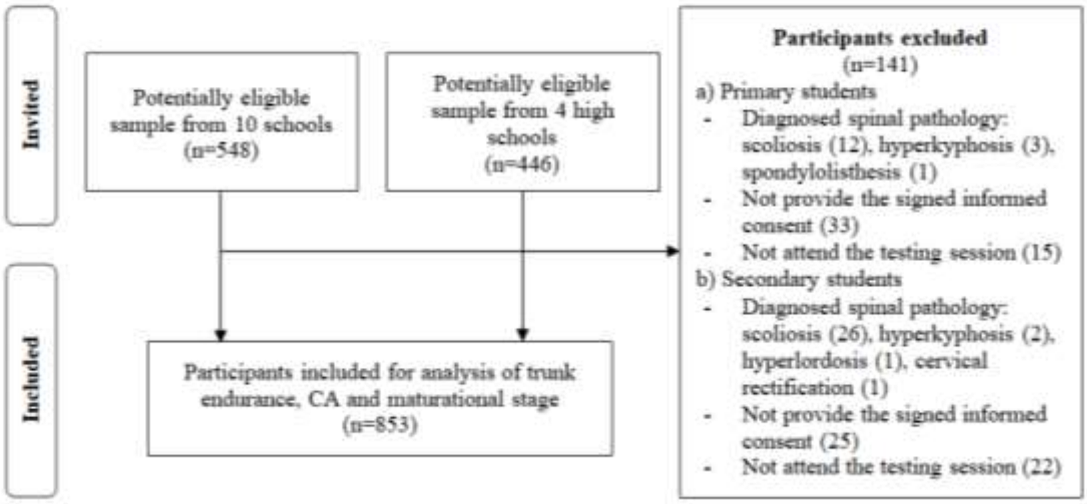


Figure 19. Flow diagram for the sample selection

Finally, a sample of 853 (age 12.18 ± 2.06 years (mean ± SD), range 9-16 years, 49% females) school and high school students completed this study. Descriptive statistics for each CA and maturation stage are displayed in Table 29, respectively. One hundred and forty-one students were removed based on the exclusion criteria (Figure 19).

Table 29. Descriptive anthropometric values (mean ± SD) for subjects per chronological age group and maturation stage

	Gender	N	Age (y)	Body mass (kg)	Body height (cm)	Leg length (cm)	Maturity offset
				Age group			
9	M	71	9.6±0.2	39.9±10	139.7±6.5	68.2±5.3	-2.7±0.3
	F	73	9.5±0.1	37.7±9	137.5±6.9	68.6±4.8	-2±0.4
10	M	79	10.5±0.2	40.3±8.3	143.5±6.3	70.1±4.5	-2.6±0.3
	F	80	10.4±0.2	41.5±11.2	142.5±6.8	68.6±4.9	-1.3±0.6
11	M	66	11.5±0.2	45.3±10.4	148.5±6.9	72.6±4.9	-2.1±0.4
	F	76	11.6±0.2	48.5±10.4	150.9±6.8	72.5±4.3	-0.1±0.4
12	M	61	12.4±0.2	51.1±12.3	154.9±8.3	75.9±5.5	-1.4±0.7
	F	56	12.4±0.2	50.8±13.1	153.4±5.8	74.2±3.4	0.3±0.5
13	M	50	13.4±0.3	56.9±13.4	162.1±8.9	79.9±5.1	-0.5±0.7
	F	34	13.4±0.3	57.4±11.7	158.1±6.5	75.6±4.2	1.3±0.4
14	M	62	14.5±0.2	63.8±13.4	167.5±6.7	82.4±4.4	0.3±0.7
	F	46	14.5±0.2	58.6±11.9	161.3±5.6	77.8±5.1	1.9±0.4
>15	M	45	15.8±0.6	67.2±15.8	172.9±8	84.9±5.2	1.5±0.7
	F	54	15.8±0.5	59.9±10.5	160±5.8	76.8±4.5	2.4±0.3
				Maturity status			
Pre-PHV	M	164	11.5±1.1	45.4±8.9	148.8±6.3	72.6±5.1	-2.1±0.6
	F	115	9.9±0.5	37±7.8	138.5±5.7	67.2±5.2	-1.8±0.5
Circa-PHV	M	72	13.9±0.9	59.9±10.6	165.2±5.7	81.1±4.9	0.02±0.4
	F	105	11.8±0.7	50±8.8	152.2±4.9	73.3±3.7	0.02±0.3
Post-PHV	M	46	15.4±0.8	73.2±14.7	175.4±5.9	85.5±4.7	1.6±0.5
	F	125	14.6±1	60.1±11.6	160.2±5.2	76.9±4.5	1.9±0.5

Note: PHV: peak height velocity, M: males, F: females, y: years, kg: kilograms, cm: centimetres

3.2.3. Procedure

Participants were tested during PE classes. Since PE teachers only have 2 sessions of 60 minutes per age-based grade per week, a time-efficient assessment procedure was designed which were divided into 2 different parts within a single testing session. The first part of each testing session was used to record the anthropometric measures needed to calculate the maturation stage of the students. The second part was designed to assess the trunk muscles endurance obtained from five field-based tests. Participants were asked not to perform strenuous exercises in the 24 h before testing sessions.

Two field-based tests were selected to assess isometric (BS test) [32] and dynamic (DEE test) [33] trunk extensor endurance, whereas 2 field-based tests were selected to assess isometric (Ito test) [34] and dynamic (BTC test) [35] trunk flexor endurance, and one field-based test to assess isometric (SB-R and SB-L test) lateral flexor endurance [36] (Figure 20). These tests were selected to measure the endurance of various areas of the trunk and so obtain a larger understanding of the overall trunk muscles.

In a previous study with high school students, the reliability of the selected field-based tests was analyzed, showing acceptable relative reliability (ICC>0.7), but absolute reliability categorized as poor (typical percentage error as a coefficient of variation from 11.3 to 24.3). These results showed that two testing sessions are not enough to make the learning effect negligible. Therefore, the PE teachers were asked to practice the different field-based tests with their students before the testing session, carrying out a minimum of 4 familiarization sessions. Thus, on the day of the assessment, each participant was familiarized with the performance of the 5 endurance field-based tests.

At the start of the testing session, all the participants received comprehensive instructions for the tests, and their questions regarding the protocols were answered. After recording anthropometric measurements, all participants completed their usual warm-up, which was led by their PE teachers and consisted of 6-10 min of low-to moderate-intensity (self-perceived) running (including forward/backward movements and side-stepping) and general mobilization (i.e., arm circles, leg kicks) followed by 4-6 min of dynamic stretching. Due to time constraints, a randomized circuit approach was used to carry out all the tests. Five different stations were set (one for each trunk endurance test). Groups of 4-6 participants were randomly assigned to one of the 5 stations and performed the 5 field-based trunk endurance tests with a 5-min rest between each test. Five researchers, who were sports science specialists with more than 5 years of experience in neuromuscular performance assessments, were supervising each field-based test.

An extendable goniometer (Lafayette Instrument Co, Lafayette, IN, USA) was used to ensure the correct joint position was maintained during the tests (BS and DEE test). During the performance of all field-based tests, participants were strongly encouraged verbally to maintain the position as long as possible or to perform the maximum number of repetitions as possible.

Anthropometry and stage of maturation

Body mass in kilograms was measured on a calibrated physician scale (SECA 799, Hamburg, Germany). Body height was recorded in centimetres on a measurement platform (SECA 799). Sitting height was measured in centimetres. Leg length was calculated as the difference between body height and sitting height.

Stage of maturation was calculated in a non-invasive manner using a regression equation comprising measures of age, body mass, body height, sitting height and leg length taken during the first part of the testing sessions [37]. Using this method, maturity offset (calculation of years from PHV) was completed (equation 1 for boys and equation 2 for girls). Due to the error in the prediction equation of approximately 6 months in the paediatric population [37], participants with a maturity offset of -0.99 to -0.51 years and

+0.51 to +0.99 years were removed from the analysis by maturational stage [9,27]. Also, participants whose maturity offset were outside −3 or +3 years, were removed from the analysis by maturational stage to maximize accuracy [27]. This approach enabled the identification of 3 different maturity groups: pre-PHV (maturity offset of less than −1), circa-PHV (maturity offset between −0.5 and +0.5), and post-PHV (maturity offset of greater than +1). Therefore, the following equations to calculate maturity offset were used:

Equation 1 (boys) $= -9.236 + 0.0002708 \times (Leg\ Length \times Sitting\ Height) - 0.001663 \times (Age \times Leg\ Length) + 0.007216 \times (Age \times Sitting\ Height) + 0.02292 \times (Weight/Height \times 100)$

Equation 2 (girls) $= -9.376 + 0.0001882 \times (Leg\ Length \times Sitting\ Height) + 0.0022 \times (Age \times Leg\ Length) + 0.005841 \times (Age \times Sitting\ Height) - 0.002658 \times (Age \times Weight) + 0.07693 \times (Weight/Height \times 100)$

Trunk extensor endurance field-based tests

Biering-Sorensen test

The isometric endurance of the trunk extensor musculature was assessed through the BS test [32]. The test started with the participant in a prone position with the lower body resting on a test bench and the anterior superior iliac spine aligned at the edge of the test bench. The lower body was attached to the test bench by 2 inextensible straps (knees and ankles). In the starting position, the upper body rested with both forearms placed on a chair. During the test, the upper body was maintained in a horizontal position with arms crossed on the chest while holding the head in a neutral position (Figure 20A). The test consisted of maintaining the trunk in the described position for as long as possible, until exhaustion, or until participants lost the correct position more than 3 times. The test duration was recorded in seconds.

Dynamic Extensor Endurance test

The dynamic endurance of the trunk extensor muscles was assessed through the DEE test [33]. Participants were located in the same position as the BS test. In the starting position, hip flexion of 45º was performed and both forearms rested on a chair. During the test, participants had to extend the trunk horizontally and then return to the initial position with arms crossed on the chest (Figure 20B). Participants were asked to carry out the maximum repetitions possible in 60 seconds. Only those repetitions that were performed correctly were counted, that is, those in which the trunk was fully extending (horizontally), and in which the head touched the chair when flexing the hip. The hip flexion during the test was controlled through a static reference (extendable goniometer).

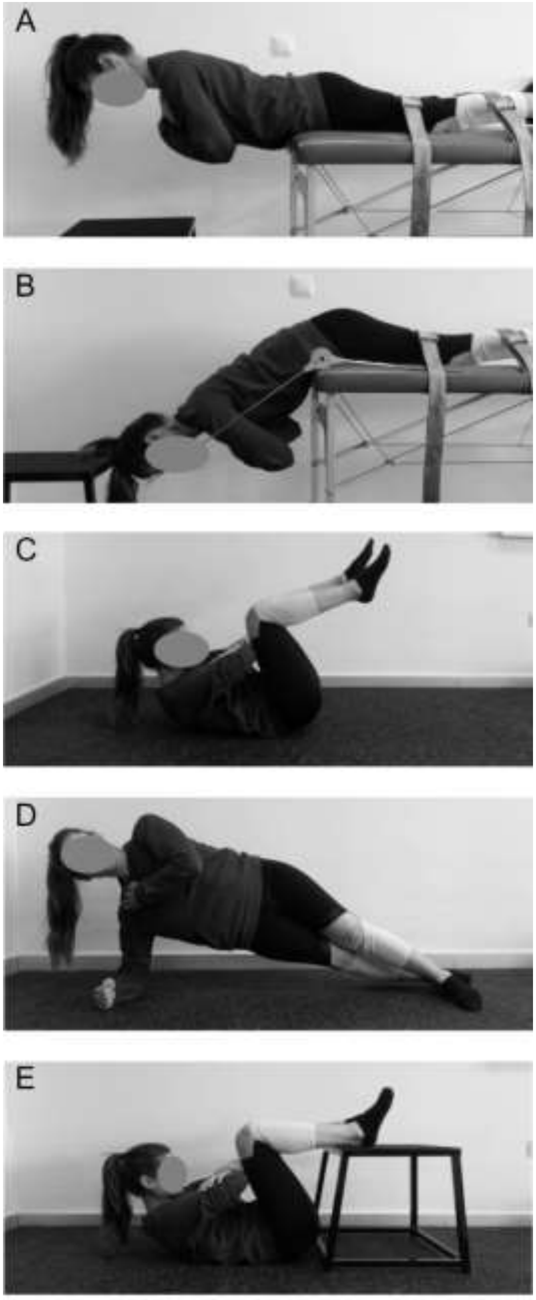


Figure 20. Trunk endurance field-based tests. A: Biering-Sorensen test, B: Dynamic Extensor Endurance test, C: Ito test, D: Side Bridge test and E: Bench Trunk Curl-Up test.

Trunk flexor endurance field-based tests

Ito test

The isometric endurance of the trunk flexor muscles was assessed through the Ito test [34]. Participants were placed in a supine position with hips and knees flexed 90° (extendable goniometer) and arms interlaced with hands grasping the opposite elbow. From this position, participants performed a trunk flexion ("curl-up") until they touched their thighs with their elbows, the scapulae did not touch the mat and the head was in a neutral position. The test consisted of maintaining this position for as long as possible, until exhaustion. The test ended when the scapulae came in contact with the mat, recording the test duration in seconds.

The original test was modified to normalize the range of motion to the participants' characteristics and thus avoid hip and lower back flexion ("sit-up") [38,39]. For this, before starting the test, the participant was placed in the aforementioned initial position, and then the subject performed a trunk flexion until the scapulae did not touch the mat. From this position, the tester approached the participant's legs towards their elbows, until they came into contact (Figure 20C). Then, the tester held the legs in this new position while the participant rested before starting the test. This leg position was maintained throughout the test.

Bench Trunk Curl-Up test

The dynamic endurance of the trunk flexor muscles was assessed through the BTC test [35]. Participants were placed in a supine position with hips and knees flexed at 90° (extendible goniometer) and resting on a bench. The arms were crossed with the hands grasping the opposite elbow (Figure 20E). From this position, participants performed a trunk flexion ("curl-up") until they touched their thighs with their elbows, the scapulae did not touch the mat and the head was in a neutral position and then returned to the initial position. Just like the Ito test, a modification of the original test was performed to avoid hip and lower back flexion ("sit-up"), approaching the participant's legs towards their elbows, until they came into contact. The test consisted of performing the maximum number of repetitions possible in 2 minutes. Only those repetitions that were performed correctly were counted, that is, those in which the elbows touched the thighs in the flexing of the trunk, and in which the head touched the mat when lowering the trunk.

Trunk lateral flexor endurance field-based test

Side Bridge test

The isometric endurance of the trunk lateral flexor musculature was assessed through the SB-R and SB-L test [36]. Participants were placed in a lateral position on their side (supported by either the dominant and non-dominant arm depending on the side tested) with legs extended. The participants were supported on their elbow and feet, the top foot was placed ahead of the lower foot (with 90º elbow flexion and the arms perpendicular to the mat) while bridging their hips off the mat to maintain an aligned body position. The uninvolved arm was held across the chest with the hand placed on the opposite shoulder (Figure 20D). The test finished when the subject lost the aligned postural position, and the duration recorded in seconds. Both sides were tested with the dominant side always examined first.

3.2.4. Statistical Analyses

Descriptive statistics including mean and SD values were calculated for each trunk endurance measure and groups separately. Not all anthropometric and trunk endurance measures were normally distributed, determined by the Kolmogorov-Smirnov test.

A Bayesian ANOVA was performed to determine the existence of between-group differences for all normal data distribution (for non-normally distributed variables a logarithmic transformation of the data were performed). In particular, separate analyses were performed to examine between-group differences for a range of CA groups (<10 years [9 to 10 years], 11 years, 12 years, 13 years, 14 years, and >15 years [15 to 16 years]). A secondary analysis was also used, grouping students by their stages of maturation (pre-PHV, circa-PHV, or post-PHV).

The BF10 was interpreted using the evidence categories suggested by Lee and Wagenmakers (2013): <1/100 = extreme evidence for H0, from 1/100 to 1/30 = very strong evidence for H0, from 1/30 to 1/10 = strong evidence for H0, from 1/10 to 1/3 = moderate evidence for H0, from 1/3 to 1 anecdotal evidence for H0, from 1 to 3 = anecdotal evidence for H1, from 3 to 10 = moderate evidence for H1, from 10 to 30 = strong evidence for H1, from 30 to 100 = very strong evidence for H1, >100 extreme evidence for H1. The median and the 95% central credible interval (CI) of the posterior distribution of the standardized effect size (δ) (i.e., the population version of Cohen's d) were also calculated for each of the between-groups comparisons. Magnitudes of the posterior distribution of the standardized effect size were classified as: trivial (<0.2), small (0.2 – 0.6), moderate (0.6 – 1.2), large (1.2 – 2.0) and very large (2.0 – 4.0) [41]. Post hoc comparisons were made to determine significant between-group differences when the fixed factor showed at least a strong (ten times higher) evidence for supporting H_1 (BF_{10} >10), an error percentage <10 and δ <0.6 (at least moderate) [40].

Finally, a Bayesian correlation coefficient analysis (Pearson's rho) was performed to examine the correlation between students' anthropometric variables and each trunk endurance score. Magnitudes of the Bayesian correlation coefficients were assessed using the thresholds defined by Hinkle, Wiersma and Jurs (2003): from 0 to 0.3 = negligible correlation, from 0.3 to 0.5 = low correlation, from 0.5 to 0.7 = moderate, from 0.7 to 0.9 high, and from 0.9 to 1 = very high correlation.

Statistical analysis was performed using the JASP computer software Version 0.11.1 (JASP Team, Amsterdam, The Netherlands).

3.3. Results

Mean values of endurance for each trunk field-based test according to CA and sex are shown in Table 30. The between-group analysis through Bayesian ANOVA showed differences with at least strong evidence in favour of the alternative hypothesis (H1) with at least moderate effect size (δ >0.6) for the endurance scores of all trunk measures obtained from BTC, Ito, DEE, BS, SB-R and SB-L tests and CA groups, regardless of sex, i.e. all trunk endurance measures increased with rising age. Furthermore, sex-by-chronological age interaction was found for BTC, DEE, SB-R and SB-L tests.

Within boys, CA differences were identified for all trunk endurance field-based tests (Table 32). Among girls, CA differences were found for all trunk endurance field-based tests, except for the DEE test where no differences were detected (Table 32). In both cases, the differences were mainly found between the youngest, worst performance, and oldest participants, highest endurance.

Differences by sex and CA group were identified for the DEE test between males and females of 14 years (BF_{10}=58.12, δ =0.99) and between those older than 15 years (BF_{10}=979, δ =1.12). Sex-related differences were also found in the SB test for the right and left side between participants older than 15 years (BF_{10}=26.61, δ =0.62; BF_{10}=45.74, δ =0.66, respectively). On the contrary, there were no significant differences found between sex and CA groups for the BTC, Ito and BS tests (Figure 21).

Table 30. Mean and SD scores obtained from trunk endurance field-based tests per chronological age group and sex.

		Trunk flexor endurance tests		Trunk extensor endurance tests		Trunk lateral flexor endurance tests	
Age group	Gender	BTC	ITO	DEE	BS	SB-R	SB-L
<10	M	46.97±22.11	63.04±42.93	31.68±13.43	78.35±42.58	34.08±24.93	38.19±28.21
	F	49.55±18.23	69.68±51.53	30.83±10.39	114.87±69.80	36.74±20.22	35.92±22.18
11	M	52.65±20.09	92.35±78.67	37.20±12.42	109.84±77.10	47.37±24.75*	50.50±27.57
	F	43.42±14.56	91.74±77.78	35.07±10.32	126.28±83.11	37.45±24.18	38.44±22.47
12	M	60.90±20.60	115.45±92.84*	41.75±10.82*	128.23±66.48*	45.75±28.25	49.28±29.52
	F	53±19.74	129.81±114.64	37.71±9.12	153.24±85.61	46.30±23.48	47.97±23.85
13	M	59.62±27.12	159.91±117.74*‡	45.16±12.47*	133.54±52.99*	45.62±27.25	51.62±26.78
	F	65.50±29.48	137±119.75**†**	37.53±12.23	155.53±62.99	57.38±32.45**†‡**	59.56±36.66**†**
14	M	61.25±19.95*	182.98±139.74*‡	50.35±12.12*‡	143.98±63.18*‡	58.36±27.43*	56.39±26.93*
	F	77.15±22.14**†‡**•	181.33±143.32**†‡**	38.18±7.94	163.35±73.77**†**	42.46±24.16	42.33±22.12
>15	M	70.52±14.99*‡	201.74±134.45*‡	51.24±11.93*‡†	160.86±62.08*‡	74.19±26.67*‡†^	70.17±24.41*‡†^
	F	62.62±16.57**†‡**	179.10±121.43**†‡**	37.47±10.59	151.72±61.09**†**	56.70±27.24**†‡**	52.89±26.35**†‡**

Note: s: seconds; rep: repetitions; BTC: Bench Trunk Curl-Up test; DEE: Dynamic Extensor Endurance test; BS: Biering-Sorensen test; SB-R: Side Bridge Right test; SB-L: Side Bridge Left test. *: significantly different from males <10 years. ‡: significantly different from males 11 years. †: significantly different from males 12 years. ^: significantly different from males 13 years. **†**: significantly different from females <10 years. **‡**: significantly different from females 11 years. •: significantly different from females 12 years.

Table 31. Mean and SD scores obtained from trunk endurance field-based tests per maturational group and sex.

		Trunk flexor endurance tests		Trunk extensor endurance tests		Trunk lateral flexor endurance tests	
Age group	Gender	BTC	ITO	DEE	BS	SB-R	SB-L
Pre-PHV	M	50.24±20.15	92.74±76.82	36.67±13.09	110.81±67.86	44.01±27.24	47.34±27.24
	F	49.53±17.94	69.26±52.22	30.97±10.58	114.78±66.37	36.86±20.26	36.02±22.46
Circa-PHV	M	66.35±20.80*	178.08±139.36*	52±10.92*	140.05±61.77	52.58±28.49*	53.87±28.19
	F	46±16.64	91.04±80.59	36±11.47	131.97±84.14	39.98±24.50	42.70±26.53
Post-PHV	M	68.58±16.34*	193.89±134.83*	47.88±12.68*	151.20±61.13*	71.04±27.05*‡	66.73±26.28*
	F	66.51±21.54**†‡**	168.95±131.13**†‡**	37.04±9.53	152.86±65.95**†**	50.53±27.31	49.29±26.03**†**

Note: PHV: peak height velocity, s: seconds; rep: repetitions; BTC: Bench Trunk Curl-Up test; DEE: Dynamic Extensor Endurance test; BS: Biering-Sorensen test; SB-R: Side Bridge Right test; SB-L: Side Bridge Left test. *: significantly different from males Pre-PHV. ‡: significantly different from males circa-PHV. **†**: significantly different from females pre-PHV. **‡**: significantly different from females circa-PHV.

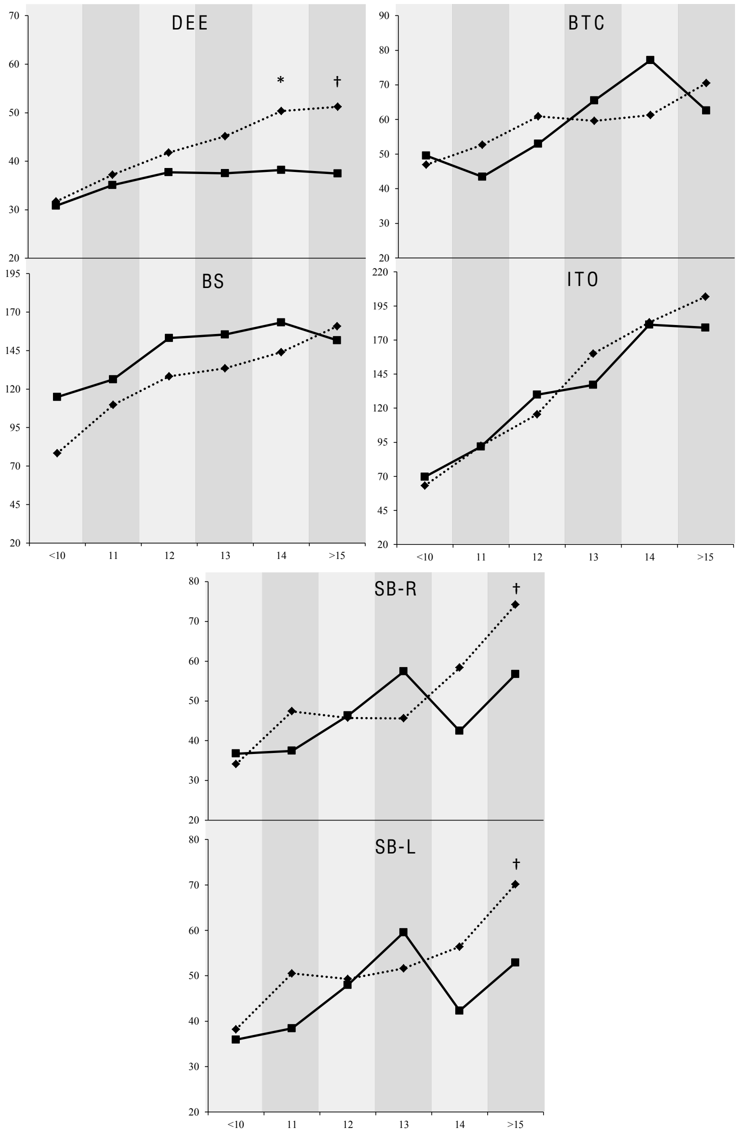


Figure 21. Mean values of endurance for each trunk field-based test according to CA and sex. ··◆·· males, ━■━ females. . *: significant differences between 14-year-old boys and girls. †: significant differences between 15-year-old boys and girls.

Mean values of endurance for each trunk field-based test according to maturity status and sex are shown in Table 31. The Bayesian ANOVA analysis showed strong evidence in favour of the H1 between maturational stages, regardless of sex, for all trunk endurance values. Furthermore, sex-by-age interaction was found for BTC, Ito and DEE tests.

Among boys, pre- and post-PHV differences were found for all trunk endurance field-based tests (Table 31). Differences between pre- and circa-PHV were only found in males for BTC, Ito, DEE and SB-R tests. Within girls, pre- and post-PHV differences were found for all trunk endurance field-based tests, except for the DEE and SB-R tests were no differences were detected (Table 33). Differences between circa- and post-PHV were only found in females for BTC and Ito tests.

For all trunk endurance field-based tests, males presented higher performance than females in circa-PHV (BTC: BF_{10}=651, δ =0.94; Ito: BF_{10}=4870, δ =0.79; DEE: BF_{10}=2,814e+6, δ =1.35), although for BS and SB tests the results are not relevant (Figure 22). Sex-related differences were also found for the DEE and SB tests between boys and girls post-growth spurt (DEE: BF_{10}=793, δ =0.97; SB-R: BF_{10}=1640, δ =0.73; SB-L: BF_{10}=95.87, δ =0.60). Sex-by-maturational stage interaction was observed for BTC, Ito and DEE tests.

Negative correlations were observed between the measures for the extensor and lateral flexor muscles tests with respect to the weight and BMI of the participants according to sex and CA. Specifically, 10- and 11-year-old boys who have a greater weight or BMI, obtain worse scores in the DEE (11-year-old, BMI (r = -0.54)) and SB right (11-year-old, weight (r = -0.54), BMI (r = -0.61)) and left (10-year-old, weight (r = -0.57), BMI (r = -0.6); 11-year-old, weight (r = -0.61), BMI (r = -0.69)) tests. On the other hand, 11- and 13-year-old girls who have a greater weight or BMI, obtain worse scores in the BS test (11-year-old, weight (r = -0.51), BMI (r = -0.6); 13-year-old, weight (r = -0.53)).

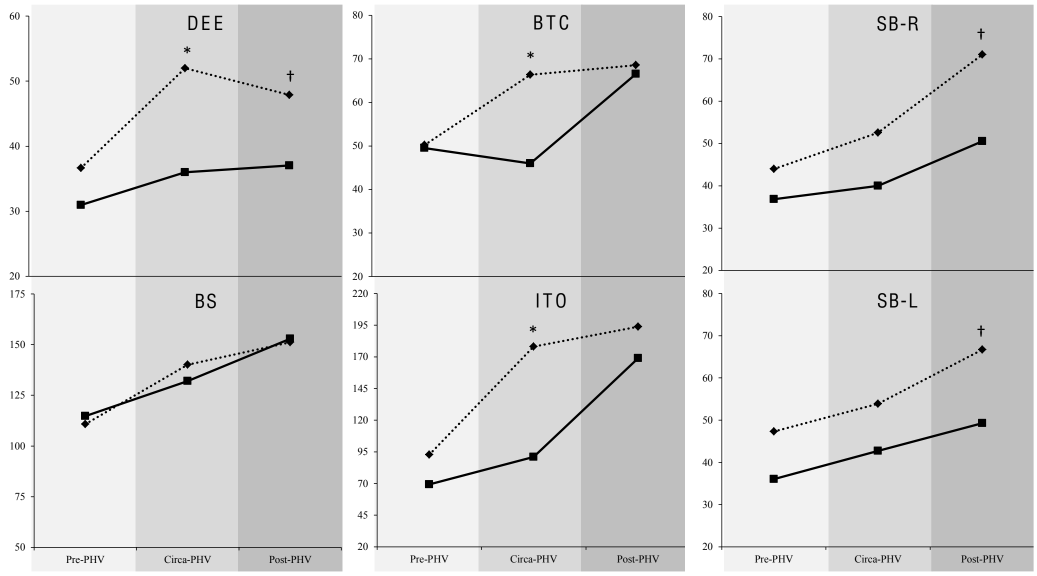


Figure 22. Mean values of endurance for each trunk field-based test according to maturity status and sex. males, females. *: significant differences between pre-PHV boys and girls. †: significant differences between post-PHV boys and girls.

Regarding the maturation groups, negative correlations were also found with weight and BMI in pre- and circa-PHV. On the one hand, the pre-PHV boys who presented greater weight and BMI showed worse values in the SB right (BMI (r = -0.51)) and left (BMI (r = -0.55)) tests. In circa-PHV boys, BTC test negatively correlated with weight (r = -0.57) and BMI (r= -0.59), BS test correlated with BMI (r = -0.51), and SB right and left tests correlated with weight (right, r= -0.60; left, r = -0.52) and BMI (right, r= -0.64;

left, r = -0.54). On the other hand, circa-PHV girls who presented greater weight (r = -0.57) and BMI (r = -0.59), showed worse values in the BS test.

3.4. Discussion

The present study aimed to analyse and compare the trunk endurance measure in children and adolescents concerning CA, maturity status and sex. The studies that have described trunk muscles endurance or strength in children and adolescents populations [13,14,28–31], have focused on the changes that occur according to CA as opposed to grouping according to maturity status. In light of existing literature, this appears to be the first study to compare trunk endurance measures across all maturity stages in a sample of both youth male and female.

The main findings of the current study showed an increase in all trunk muscle endurance measures from the younger participants (before PHV) to their older peers (after PHV) independently of the sex (i.e. all trunk endurance measures increased with the growth of the participants). Sex-related differences when analysing the trunk endurance measurements with maturity status were found, males presented higher performance than females in circa- and post-PHV.

The present study showed that the trunk muscular endurance is determined by the CA and the biological maturity of the participants. The greatest differences were found between the youngest (pre-PHV) and the oldest participants (post-PHV) for all trunk endurance measures, so these differences in performance might be influenced by physical development experienced at these ages (anatomical and hormonal changes, and changes related to motor control and physical abilities). The current results coincide with other studies carried out with children and adolescents where trunk muscle endurance or other physical demands related to performance were assessed and an increase was observed with age [13,14,26,29,30,43,44]. Furthermore, previous studies have shown that the ability to produce force is lower in children compared to adolescents or adults, partly as a result of architecture and size of the muscle [27] and specific patterns of muscle activation [45].

On the contrary, there are some researches where no differences in performance were found according to CA [28,31]. Taking into account that the greatest differences in performance usually appear between the pre- and post-PHV participants, it is reasonable that previous authors found no age-related differences in performance partly because their samples were assessment just after the PHV, Moya-Ramón et al. assessed boys and girls adolescents (mean age: 16.26 ± 1.13 years), and Papadopoulou et al. evaluated female volleyball players (mean age: 13.9 ± 1.9 years). The large inter-individual differences (magnitude of change, rate of change and onset of change) in maturity status that is often found in adolescents within the same age could influence the findings of the aforementioned studies because they can be confounding factors [27].

Although no sex-related differences were found according to CA in trunk endurance measures (except for the DEE and SB tests from 14 years), notable sex-related differences were found when analysing to the stage of maturation (Figure 3). Concerning pre-PHV, no sex-related differences were found in trunk muscle endurance, nor differences in strength or power in upper and lower extremities were found in other studies [26,44]. As mentioned above, this could be because before the growth spurt, the anatomical or hormonal changes that can cause sex-related maturation differences have not yet occurred [8,24,26].

With the onset of puberty, males showed more endurance in trunk muscles than girls, although for BS and SB tests the sex-related differences were not relevant. Sex-related differences were also found for the DEE and SB tests between boys and girls post-PHV. Previous research has found different responses in estradiol and testosterone levels in males and females during puberty [46], and this divergence could contribute to differences in fitness development between the sexes that become increasingly apparent during the puberty years. For example, increased testosterone advocates increased height and muscle mass in males, and increased estradiol promotes increased body fat and stimulates ovulation and breast development in females [14,26,47]. Therefore, the sex-related differences in trunk endurance measures found during puberty could be because in this period boys are taller, heavier, and have more fat-free mass than girls, which would result in physical performance advantage for males [8,25,28,47,48].

Given that improvements in physical fitness during development are largely influenced by changes related to growth and maturation, children must be physically active [46]. Without the stimulus of regular exercise, physical fitness gains may plateau or even decrease, especially in girls who during the growth spurt experience physical changes (i.e., widening of the hips, breast development, increased proportional fat mass) along with low proficiency in motor skills generating a negative impact on their participation in physical activities and sports [8,25,48–50]. In fact, early maturation in girls was associated with a decrease in total motor competence, which could be related to poor performance in tasks that involve weight-bearing or endurance, as the case in the field-based tests carried out in the current study [25,50].

On the other hand, differences have also been found in the development of trunk endurance through the stages of maturation according to sex. Whereas large increases in trunk endurance are found in boys when moving from pre- to circa-PHV (for BTC, Ito, DEE, and SB-R tests), the rise in women is less pronounced and more stable across the different stages of maturation. A possible explanation may be because, with the onset of puberty, increases in insulin-like growth factor 1 hormone (IGF-1) and growth hormone foment protein synthesis in both males and females, but the additional anabolic effect of increased testosterone means that boys experience much greater gains in mass muscle during puberty than girls [46]. This phenomenon together with the increase in the prevalence of back pain after age 12 [1,5–7], especially among girls, suggests that the disparity found in trunk endurance muscles, both flexor and extensor muscles endurance was significantly higher among boys [13,26,43], could be one of the possible causes of the development and increase of these back troubles in females during the onset of puberty.

Finally, the negative correlations found between the measurements of the extensor and lateral flexor muscle tests to weight and BMI of the participants according to CA, maturational stage and sex support the results found in the literature. There are numerous studies where it is observed that children and adolescents had more difficulty performing the tests than involved moving your body mass or holding it in position as the body weight increases [8,25,26,31,51,52]. For example, in the study of Ervin et al. (2014), muscular endurance was assessed in boys and girls aged 6 to 15 years and they observed that the performances on the plank and modified pull-up were lower for overweight and obese youths compared with normal-weight youth, regardless of sex. Papadopoulou et al. (2020), when assessing the trunk endurance in volleyball players, also found negative correlations between the percentage of body fat and the performance in the tests. That is, having a higher body weight or BMI, before and during the pubertal stretch, can negatively influence the performance of field tests for trunk muscular endurance. Given that during the tests for extensors and lateral flexors, the participants must maintain a certain position against gravity, the mass of the participants could make it difficult to carry out the test and negatively influence the results obtained in it. Therefore, and taking into account the different changes that occur during the growth stages, it is important for youth fitness specialists to monitor physical changes longitudinally (body mass, lean mass, fat mass, height, limb length and ratios) because it can help to partly distinguish gains due to training effects over and above those experienced simply due to growth and maturation [46].

There are some limitations to consider when interpreting the data from this study. First, the design of the cross-sectional study does not allow for the measurement of muscle changes in individual children and adolescents over time. Second, the results were obtained and were limited to Spanish children and adolescents, who may or may not represent students of the same age throughout the world. Third, the differences in the number of participants between the groups and the large variability in test results (i.e., the Ito test) could have affected the comparison between the groups. Furthermore, special attention would be required when performing trunk endurance field-based tests such as trunk flexors, as even a minimal deviation from correct position would result in altered muscle activation influencing the outcome [53]. Maturity offset was calculated using an equation based on the subjects' leg length, sitting height, age, body height, and body weight, which may not be as accurate as using skeletal images. However, to minimize the group allocation error derived from the equation, participants with a maturational offset between -1 to-0.5 and +0.5 to +1 were removed from the data set, as were the students who were outside -3 or +3 years. Subsequently, this decision led to smaller sample size in the circa-PHV group compared to the other groups. However, the large total sample size attempted to mitigate differences in the group's sample size distribution. Finally, there was the possible influence of personal factors, motivation, for example, that may influence the performance of field tests, even though participants received complete information about the tests, along with encouragement during the execution.

On the other hand, the strength of this study was the use of field-based tests to assess the trunk endurance, both isometric and dynamic, as well as in the sagittal and frontal planes, encompassing all the muscles that participate in trunk movements. Also, this study is one of few that assess participants of both sexes and from different maturational stages.

Future applied studies should carry out a longitudinal study approach to help to increase the knowledge about biological maturity in association with trunk muscles endurance, in addition to including other variables such as the alignment of the spine, the classification of the participants as active or sedentary, or according to participants with normal weight, overweight or obesity.

3.5. Conclusion

The current study provides a cross-sectional analysis of five trunk endurance field-based tests in a large sample of children and adolescents of both sexes. The findings may assist practitioners by providing a clearer understanding of expected trends in the development of trunk muscular endurance during student growth and maturation. There is a clear trend towards increasing trunk endurance as the participants are older or more mature. On the other hand, no sex-related performance differences were found according to chronological age, but sex-related differences were found according to maturational status in most of the endurance tests, this does not support the stratification of trunk endurance test scores by age used in international batteries of youth physical fitness tests.

In conclusion, pre-pubertal boys and girls showed similar performance in all field-based tests, when they entered the growth spurt the differences were relevant (BTC, Ito, DEE), with boys showing higher scores than girls in all tests. After the PHV, the differences were maintained for the lateral trunk flexors (SB test) and the DEE test and were equalized for the trunk flexor tests (BTC and Ito tests) and the BS test.

Therefore, as a result of the findings obtained, it is important to consider biological maturation, rather than chronological age, for the global exercise prescription for school-age youth, and for the development of appropriate training programs (depending on the unique physical and physiological processes that occur as a result of maturation) that optimize adaptation to training and minimize the risk of activity-related injuries, and include preventive intervention strategies to support balanced development of trunk muscular endurance during growth, particularly during the start of the growth spurt.

3.6. References

1. Beynon, A.M.; Hebert, J.J.; Lebouef-Yde, C.; Walker, B.F. Potential risk factors and triggers for back pain in children and young adults. A scoping review, part I: Incident and episodic back pain. *Chiropr. Man. Ther.* 2019, *27*(58), 1–7, doi: 10.1186/s12998-019-0280-9.

2. Junge, T.; Wedderkopp, N.; Boyle, E.; Kjaer, P. The natural course of low back pain from childhood to young adulthood - a systematic review. *Chiropr. Man. Ther.* 2019, *27*(1), 1–10, doi: 10.1186/s12998-018-0231-x.

3. Kędra, A.; Kolwicz-Gańko, A.; Sitarski, D.; Kędra, P.; Czaprowski, D.; Kedra, A.; et al. Prevalence of back pain and the knowledge of preventive measures in a cohort of 11619 Polish school-age children and youth—an epidemiological study. *Medicine (Baltimore).* 2019, *98*(22), 1–6, doi: 10.1097/MD.0000000000015729.

4. Meziat Filho, N.; Coutinho, E.S.; Azevedo e Silva, G. Association between home posture habits and low back pain in high school adolescents. *Eur. Spine J.* 2015, *24*(3), 425–33, doi: 10.1097/MD.0000000000015729.

5. Hebert, J.J.; Leboeuf-Yde, C.; Franz, C.; Lardon, A.; Hestbæk, L.; Manson, N.; et al. Pubertal development and growth are prospectively associated with spinal pain in young people (CHAMPS study-DK). *Eur. Spine J.* 2019, *28*(7), 1565–71, doi: 10.1007/s00586-019-05905-6.

6. Lardon, A.; Leboeuf-Yde, C.; Le Scanff, C.; Wedderkopp, N. Is puberty a risk factor for back pain in the young? A systematic critical literature review. *Chiropr. Man. Ther.* 2014, *22*(1), 1–12, doi: 10.1186/s12998-014-0027-6.

7. Calvo-Muñoz, I.; Gómez-Conesa, A.; Sánchez-Meca, J. Prevalence of low back pain in children and adolescents: A meta-analysis. *BMC Pediatr.* 2013, *13*(14), 1–12, doi: 10.1186/1471-2431-13-14.

8. Malina, R.M.; Bouchard, C.; Bar-Or, O. *Growth, Maturation, and Physical Activity*. 2nd ed. Human Kinetics: Champaign, USA, 2004.

9. Robles-Palazón, F.J.; Ayala, F.; Cejudo, A.; De Ste Croix, M.; Sainz de Baranda, P.; Santonja, F. Effects of Age and Maturation on Lower Extremity Range of Motion in Male Youth Soccer Players. *J. Strength Cond. Res.* 2020, *Publish Ah*, doi: 10.1519/jsc.0000000000003642.

10. Lloyd, R.S.; Oliver, J.L.; Faigenbaum, A.D.; Myer, G.D.; De Ste Croix, M.B. Chronological age vs. biological maturation: implications for exercise programming in youth. *J. Strength Cond. Res.* 2014, *28*(5), 1454–64, doi: 10.1519/JSC.0000000000000391.

11. John, C.; Rahlf, A.L.; Hamacher, D.; Zech, A. Influence of biological maturity on static and dynamic postural control among male youth soccer players. *Gait Posture.* 2019, *68*, 18–22, doi: 10.1016/j.gaitpost.2018.10.036.

12. MacDonald, J.; Stuart, E.; Rodenberg, R. Musculoskeletal low back pain in school-aged children a review. *JAMA Pediatr.* 2017, *171*(3), 280–7, doi: 10.1001/jamapediatrics.2016.3334.

13. Dejanovic, A.; Harvey, E.P.; McGill, S.M. Changes in Torso Muscle Endurance Profiles in Children Aged 7 to 14 Years: Reference Values. *Arch. Phys. Med.*

Rehabil. 2012, *93*(12), 2295–301, doi: 10.1016/j.apmr.2011.12.023.

14. Dejanovic, A.; Cambridge, E.D.J.; McGill, S. Isometric torso muscle endurance profiles in adolescents aged 15-18: normative values for age and gender differences. *Ann. Hum. Biol.* 2014, *41*(2), 153–8, doi: 10.3109/03014460.2013.837508.

15. Czaprowski, D.; Stoliński, Ł.; Tyrakowski, M.; Kozinoga, M.; Kotwicki, T. Non-structural misalignments of body posture in the sagittal plane. *Scolisis Spinal Disord.* 2018, *13*(6), 1–14, doi: 10.1186/s13013-018-0151-5.

16. Lardon, A.; Leboeuf-Yde, C.; Le Scanff, C. Is back pain during childhood or adolescence associated with muscle strength, muscle endurance or aerobic capacity: Three systematic literature reviews with one meta-analysis. *Chiropr. Man. Ther.* 2015, *23*(1), 1–12, doi: 10.1186/s12998-015-0065-8.

17. Potthoff, T.; De Bruin, E.D.; Rosser, S.; Humphreys, B.K.; Wirth, B. A systematic review on quantifiable physical risk factors for non-specific adolescent low back pain. *J. Pediatr. Rehabil. Med.* 2018, *11*(2), 79–94, doi: 10.3233/PRM-170526.

18. Abdelraouf, O.R.; Abdel-aziem, A.A. The relationship between core endurance and back dysfunction in collegiate male athletes with and without nonspecific low back pain. *Int. J. Sports Phys. Ther.* 2016, *11*(3), 337–44.

19. Johnson, O.; Mbada, C.; Akosile, C.; Agbeja, O. Isometric endurance of the back extensors in school-aged adolescents with and without low back pain. *J. Back Musculoskelet. Rehabil.* 2009, *22*(4), 205–11, doi: 10.3233/BMR-2009-0235.

20. Astfalck, R.G.; O'Sullivan, P.B.; Straker, L.M.; Smith, A.J. A detailed characterisation of pain, disability, physical and psychological features of a small group of adolescents with non-specific chronic low back pain. *Man. Ther.* 2010, *15*(3), 240–7, doi: 10.1016/j.math.2009.12.007.

21. Calvo-Muñoz, I.; Gomez-Conesa, A.; Sanchez-Meca, J. Preventive physiotherapy interventions for back care in children and adolescents: a meta-analysis. *BMC Musculoskelet. Disord.* 2012, *13*(152), 1–19, doi: 10.1186/1471-2474-13-152.

22. Steffens, D.; Maher, C.G.; Pereira, L.S.M.; Stevens, M.L.; Oliveira, V.C.;

Chapple, M.; et al. Prevention of low back pain. A systematic review and meta-analysis. *JAMA Intern. Med.* 2016, *176*(2), 199–208, doi: 10.1001/jamainternmed.2015.7431.

23. De Ste Croix, M.B.A.; Deighan, M.A.; Armstrong, N. Assessment and interpretation of isokinetic muscle strength during growth and maturation. *Sport Med.* 2003, *33*(10), 727–43, doi: 10.2165/00007256-200333100-00002.

24. Shultz, S.J.; Nguyen, A.D.; Schmitz, R.J. Differences in lower extremity anatomical and postural characteristics in males and females between maturation groups. *J. Orthop. Sports Phys. Ther.* 2008, *38*(3), 137–49, doi: 10.2519/jospt.2008.2645.

25. Cumming, S.P.; Standage, M.; Gillison, F.; Malina, R.M. Sex Differences in Exercise Behavior During Adolescence: Is Biological Maturation a Confounding Factor? *J. Adolesc. Heal.* 2008, *42*(5), 480–5, doi: 10.1016/j.jadohealth.2007.10.005.

26. Ervin, R.; Fryar, C.; Wang, C-Y.; Miller, I.; Ogden, C. Strength and body weight in US children and adolescents. *Pediatrics.* 2014, *134*(3), e782–9, doi: 10.1542/peds.2014-0794.

27. Radnor, J.M.; Oliver, J.L.; Waugh, C.M.; Myer, G.D.; Lloyd, R.S. The influence of maturity status on muscle architecture in school-aged boys. *Pediatr. Exerc. Sci.* 2020, *32*(2), 89–96, doi: 10.1123/pes.2019-0201.

28. Moya-Ramón, M.; Juan-Recio, C.; Lopez-Plaza, D.; Vera-Garcia, F.J. Dynamic trunk muscle endurance profile in adolescents aged 14-18: Normative values for age and gender differences. *J. Back Musculoskelet. Rehabil.* 2018, *31*(1), 155–62, doi: 10.3233/BMR-169760.

29. Arampatzis, A.; Frank, J.; Laube, G.; Mersmann, F. Trunk muscle strength and lumbo-pelvic kinematics in adolescent athletes: Effects of age and sex. *Scand. J. Med. Sci. Sport.* 2019, *29*(11), 1691–8, doi: 10.1111/sms.13503.

30. Mueller, J.; Mueller, S.; Stoll, J.; Baur, H.; Mayer, F. Trunk extensor and flexor strength capacity in healthy young elite athletes aged 11-15 years. *J. Strength Cond. Res.* 2014, *28*(5), 1328–34, doi: 10.1519/JSC.0000000000000280.

31. Papadopoulou, S.D.; Zorzou, A.; Drikos, S.; Stavropoulos, N.; Knechtle, B.;

Nikolaidis, P.T. Core stability and symmetry of youth female volleyball players: A pilot study on anthropometric and physiological correlates. *Symmetry (Basel).* 2020, *12*(2), doi: 10.3390/sym12020249.

32. Biering-Sorensen, F. Physical Measurements as Risk Indicators of Low-Back trouble over one-year period. *Spine (Phila Pa 1976).* 1984, *9*(2), 106–19.

33. Luoto, S.; Heliövaara, M.; Hurri, H.; Alaranta, H. Static back endurance low-back pain and the risk of low-back pain. *Clin. Biomech.* 1995, *10*(6), 323–4, doi: 10.1016/0268-0033(95)00002-3.

34. Ito, T.; Shirado, O.; Suzuki, H.; Takahashi, M.; Kaneda, K.; Strax, T.E. Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. *Arch. Phys. Med. Rehabil.* 1996, *77*, 75–9, doi: 10.1016/S0003-9993(96)90224-5.

35. Knudson, D.; Johnston, D. Validity and Reliability of a Bench Trunk-Curl Test of Abdominal Endurance. *J. Strength Cond. Res.* 1995, *9*(3), 165–9.

36. McGill, S.M.; Childs, A.; Liebenson, C. Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. *Arch. Phys. Med. Rehabil.* 1999, *80*(8), 941–4, doi: 10.1016/S0003-9993(99)90087-4.

37. Mirwald, R.L.; Baxter-Jones, A.D.G.; Bailey, D.A.; Beunen, G.P. An assessment of maturity from anthropometric measurements. *Med. Sci. Sports Exerc.* 2002, *34*(4), 689–94, doi: 10.1097/00005768-200204000-00020.

38. Juan-Recio, C.; Barbado, D.; López-Valenciano, A.; Vera-García, F.J. Test de campo para valorar la resistencia de los músculos del tronco. *Apunt Educ. Fis. y Deport.* 2014, *3*(117), 59–68, doi: 10.5672/apunts.2014-0983.es.(2014/3).117.06.

39. Juan-Recio, C.; Lopez-Vivancos, A.; Moya, M.; Sarabia, J.M.; Vera-García, F.J. Short-term effect of crunch exercise frequency on abdominal muscle endurance. *J. Sports Med. Phys. Fitness.* 2015, *55*(4), 280–9,

40. Lee, M.D.; Wagenmakers, E.J. *Bayesian cognitive modeling: A practical course.* Cambirdge University Press: United Kingdom, 2013.

41. Batterham, A.M.; Hopkins, W.G. Making meaningful inferences about magnitudes. *Int. J. Sports Physiol. Perform.* 2006, *1*, 50–7, doi:

10.1123/ijspp.1.1.50.

42. Hinkle, D.; Wiersma, W.; Jurs, S. *Applied Statistics for the Behavioral Sciences*. 5th ed. Houghton Mifflin: Boston, USA, 2003.

43. Bernard, J-C.C.; Bard, R.; Pujol, A.; Combey, A.; Boussard, D.; Begue, C.; et al. Muscle assessment in healthy teenagers. Comparison with teenagers with low back pain. *Ann. Readapt. Med. Phys.* 2008, *51*(4), 274–83, doi: 10.1016/j.annrmp.2008.03.010.

44. Fry, A.C.; Irwin, C.C.; Nicoll, J.X.; Ferebee, D.E. Muscular strength and power in 3- To 7-year-old children. *Pediatr. Exerc. Sci.* 2015, *27*(3), 345–54, doi: 10.1123/pes.2014-0152.

45. Dotan, R.; Mitchell, C.; Cohen, R.; Klentrou, P.; Gabriel, D.; Falk, B. Child-Adult Differences in Muscle Activation-A Review. *Pediatr. Exerc. Sci.* 2012, *24*(1), 2–21.

46. Faigenbaum, A.D.; Lloyd, R.S.; Oliver, J.L. *Essentials of Youth Fitness*. Human Kinetics: Champaign, USA, 2019.

47. Round, J.M.; Jones, D.A.; Honour, J.W.; Nevill, A.M. Hormonal factors in the development of differences in strength between boys and girls during adolescence: A longitudinal study. *Ann. Hum. Biol.* 1999, *26*(1), 49–62, doi: 10.1080/030144699282976.

48. Malina, R.M.; Rogol, A.D.; Cumming, S.P.; Coelho E Silva, M.J.; Figueiredo, A.J. Biological maturation of youth athletes: Assessment and implications. *Br. J. Sports Med.* 2015, *49*(13), 852–9, doi: 10.1136/bjsports-2015-094623.

49. Garcia, C.; Teles, J.; Barrigas, C.; Fragoso, I. Health-related quality of life of Portuguese children and adolescents according to their biological maturation and volume of physical activity. *Qual. Life Res.* 2018, *27*(6), 1483–92, doi: 10.1007/s11136-018-1822-7.

50. Drenowatz, C.; Greier, K. Association of biological maturation with the development of motor competence in Austrian middle school students - A 3-year observational study. *Transl. Pediatr.* 2019, *8*(5), 402–11, doi: 10.21037/TP.2019.09.03.

51. Dejanovic, A.; Harvey, E.; Andersen, J.; Mcgill, S.M. Do Anthropometric

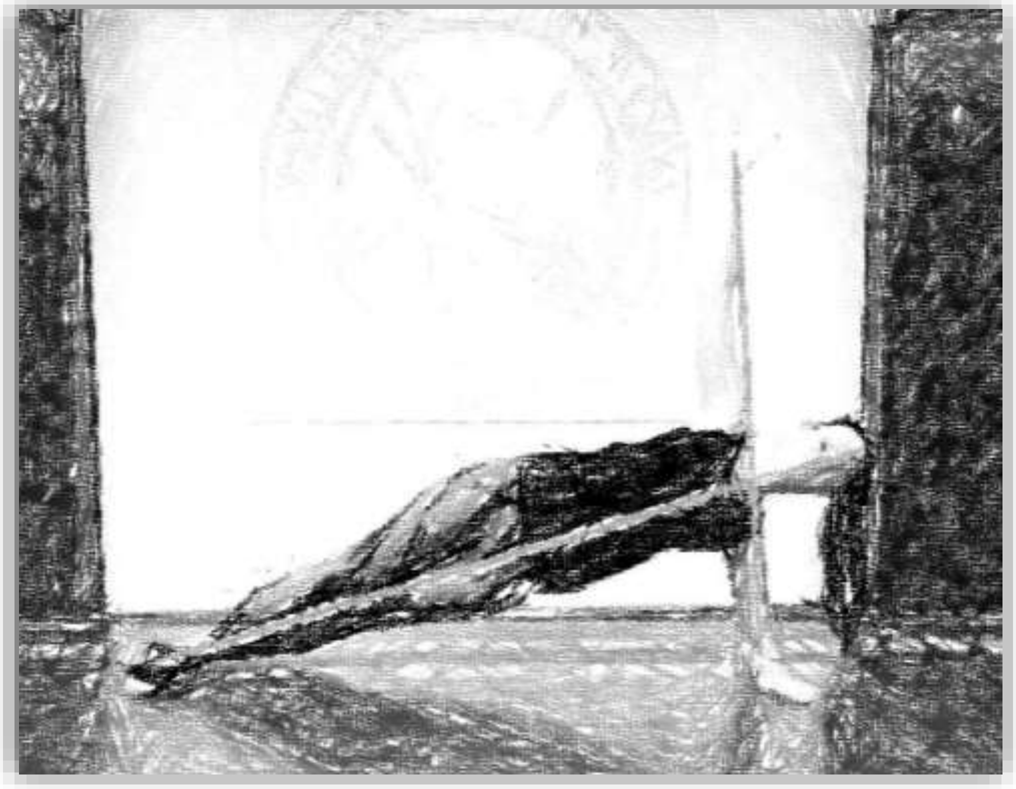
Measures Influence Torso Muscle Endurance Profiles of Children Aged 7 to 14? *Adv. Phys. Educ.* 2012, *2*(4), 187–96, doi: 10.1016/j.apmr.2011.12.023.

52. Hulteen, R.M.; Morgan, P.J.; Barnett, L.M.; Stodden, D.F.; Lubans, D.R. Development of Foundational Movement Skills: A Conceptual Model for Physical Activity Across the Lifespan. *Sport Med.* 2018, *48*(7), 1533–40, doi: 10.1007/s40279-018-0892-6.

53. Tse, M.A.; McManus, A.M.; Masters, R.S. Trunk muscle endurance tests: effect of trunk posture on test outcome. *J. Strength Cond. Res.* 2010, *24*(12), 3464–70, doi: 10.1519/JSC.0b013e3181aeb195.

CAPÍTULO IV

APLICACIONES PRÁCTICAS



4. Aplicaciones prácticas

En base a los resultados obtenidos en los diferentes estudios que forman la presente tesis doctoral, se pueden extraer las siguientes aplicaciones prácticas:

Línea de investigación 1:

- La prevalencia de dolor de espalda entre niños y adolescentes de la Región de Murcia presenta diferentes patrones en función del sexo.

En el presente estudio se ha encontrado que, en las chicas, la aparición del dolor de espalda está asociado con el desarrollo puberal y en los chicos, con el exceso de peso. Por lo que, a la hora de llevar a cabo programas de intervención con el fin de reducir el riesgo de sufrir dolor de espalda entre los escolares, se deberán de plantear de forma diferente en función del sexo y teniendo en cuenta el estado madurativo para así conseguir una prevención más efectiva. Los ejercicios que se realizan con el Programa ISQUIOS podrían programarse en función de la maduración de los participantes debido a que las chicas inician el proceso de maduración una media de dos años antes que los chicos, y esa etapa es determinante para la aparición del dolor de espalda.

- Los factores físicos o biomecánicos asociados al dolor de espalda en niños y adolescentes de la Región de Murcia también presentan diferentes patrones en función del sexo.

En el presente estudio se ha valorado la disposición sagital de la columna vertebral, el rango de movimiento de la musculatura de la cadera y la resistencia de la musculatura del tronco y se ha visto que su asociación con el dolor de espalda es diferente según se trate de niños o niñas. De forma general, el dolor de espalda se asoció con un exceso de peso, una reducción del rango de movimiento en la musculatura flexora de la cadera y debilidad en la musculatura flexora-lateral del tronco. A raíz de los resultados de la presente tesis, junto con los ejercicios que ya se llevan a cabo en el Programa ISQUIOS, se deberían de incluir ejercicios para mejorar también la extensibilidad de los flexores de cadera, sobre todo en las chicas, e incluir más ejercicios en el plano frontal para la mejora de la musculatura flexora-lateral, sin descuidar el trabajo para la musculatura isquiosural y del resto de musculatura del tronco.

Línea de investigación 2:

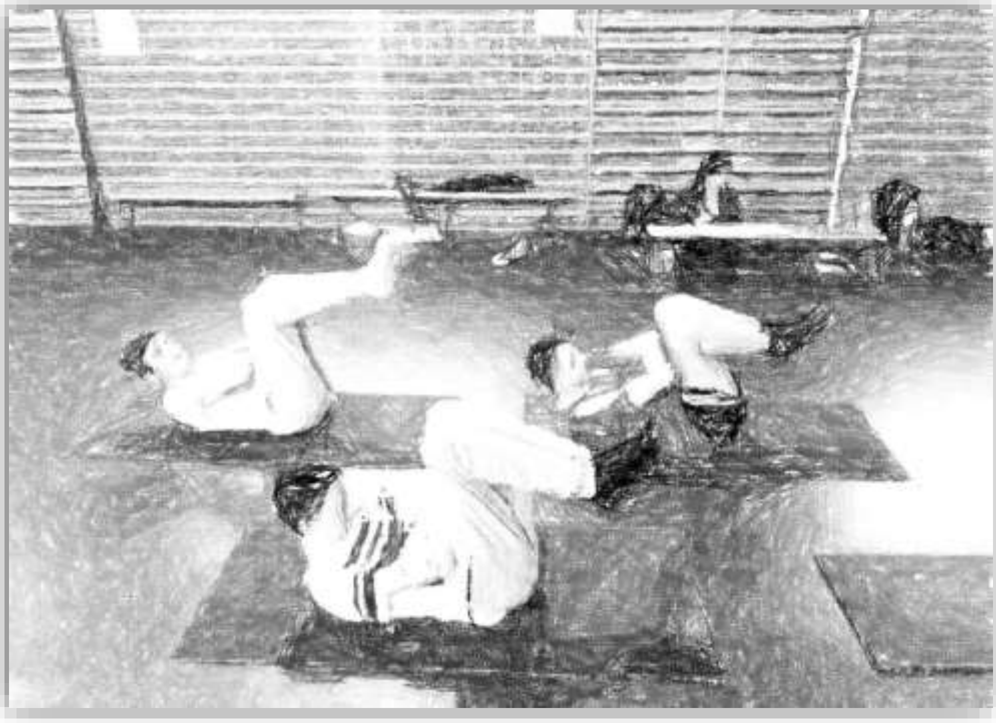
- La valoración global de la resistencia de la musculatura del tronco debería de realizarse tanto en centros educativos como en el ámbito deportivo para la realización de programas de prevención, intervención o entrenamiento de forma óptima.

Teniendo en cuenta que los déficits en la resistencia de la musculatura del tronco se han visto asociados con el dolor de espalda en niños y adolescentes, sobre todo en la musculatura flexora-lateral, es importante llevar a cabo valoraciones correctas de dicha musculatura. Como no existe una única prueba que valore todos los músculos del tronco, se deberían llevar a cabo una por cada grupo muscular para obtener una visión más global de cómo están y en función del objetivo, se podrán realizar pruebas isométricas o dinámicas. Sin embargo, es necesario seguir estudiando la fiabilidad de dichas pruebas en población juvenil para determinar las consideraciones que se deben de tener en cuenta para su correcto uso.

De forma general, animamos a que aquellos profesionales que trabajen con niños o adolescentes a seguir trabajando para la mejora de la salud postural de los jóvenes, así como a seguir formándose y actualizándose para ofrecer lo mejor a sus alumnos.

CAPÍTULO V

CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN



5. 1. CONCLUSIONES

A continuación, se presentan las principales conclusiones obtenidas a partir de los estudios que forman la presente tesis doctoral:

Línea de investigación 1:

Estudio 1: El presente estudio mostró una prevalencia de DE asociado al estado de maduración y al peso de los participantes, encontrándose diferentes patrones de prevalencia según el sexo. Tanto el DE como el DL en los niños se asoció con sobrepeso u obesidad, mientras que en las niñas se asoció con un mayor desarrollo puberal. Las características del DE también se asociaron con el peso y el estado de maduración, independientemente del sexo, encontrando una mayor frecuencia y limitación de las AVD entre los participantes más desarrollados o con sobrepeso / obesidad. Por tanto, el presente estudio respalda la afirmación de que la prevalencia del DE debe ser informada por zonas, sexo y etapa de maduración, y complementada con sus características, frecuencia e intensidad como mínimo, para ofrecer una visión completa de la evolución y severidad del DE en población juvenil.

El peso es un factor de riesgo modificable cuya optimización podría generar beneficios en la salud músculo-esquelética de los escolares y, si bien el sexo, la edad y el estado de madurez son factores biológicos intransitables, todos deben ser tomados en cuenta a la hora de aplicar programas de prevención educativos.

Estudio 2: El presente estudio ha observado que los niños y adolescentes pre-púberes o que han iniciado su desarrollo puberal, presentan diferencias relacionadas con el sexo en la evaluación del protocolo “Fitness Postural”. Los chicos presentaron un morfotipo más “hipercifótico” junto con una mayor tendencia hacia la retroversión pélvica y las chicas un morfotipo más “hiperlordótico” y una posición pélvica más neutra. Por otro lado, los chicos tenían un ROM de cadera reducido y más resistencia en los músculos flexores-laterales del tronco. Las chicas, sin embargo, mostraron más ROM en la cadera y más resistencia de los músculos extensores del tronco.

En cuanto al DE (dolor de espalda), se encontró asociación entre el DE en los chicos y el IMC más alto, y entre el DE en las chicas y un ROM reducido en los músculos flexores de la cadera, junto con una baja resistencia de los flexores del tronco. Al analizar las asociaciones entre el protocolo "Fitness Postural" y el DL, los chicos con DL mostraron asociación con un mayor IMC y una peor resistencia de los flexores-laterales del tronco, y las chicas con DL mostraron asociación con un ROM reducido en los músculos flexores de la cadera y una peor resistencia para músculos flexores del tronco y flexores-laterales.

El ROM de la cadera y la resistencia de la musculatura del tronco son factores de riesgo modificables cuya optimización puede generar beneficios en la salud músculo-esquelética de niños y adolescentes, teniendo en cuenta siempre su sexo y madurez biológica, así como la alineación de la columna, a la hora de aplicar programas de prevención educativos.

Línea de investigación 2:

Estudio 3: Los principales hallazgos del meta-análisis realizado de generalización de la fiabilidad informan que, solo la medida de resistencia de extensión del tronco obtenida mediante la prueba de Biering-Sorensen puede presentar evidencia científica suficiente en términos de fiabilidad inter- (ICC + = 0,94, basado en 8 cohortes) e intra-examinador (intra-sesión [ICC + = 0,88, basado en 12 cohortes] e inter-sesión [ICC + = 0,88, basado en 27 cohortes]) para justificar su uso (principalmente en varones adultos) con fines prácticos y clínicos. Además, esta medida de resistencia de extensión del tronco puede ser incluso más fiable cuando se lleva a cabo previamente una sesión de familiarización con el procedimiento de prueba de Biering-Sorensen y se utiliza una plomada para controlar la posición del participante durante la prueba.

Ninguna de las demás medidas de resistencia de extensión del tronco obtenidas a través de las otras pruebas de campo estáticas y dinámicas, mostraron evidencia científica suficiente (en términos de fiabilidad) para promover su utilización con fines prácticos y clínicos. Por lo tanto, se necesitan más estudios de fiabilidad en diferentes cohortes antes de que se pueda recomendar el uso de estas pruebas de campo para evaluar la resistencia del extensor del tronco en entornos prácticos y de investigación.

Finalmente, la alta tasa de inducción de la fiabilidad encontrada en el presente estudio (72,8%), sugiere que los investigadores deberían ser más conscientes de que la inducción de la fiabilidad es una práctica errónea que debe ser erradicada porque puede provocar errores en la estimación de las medidas utilizadas. Por lo tanto, en investigaciones futuras, los autores deben informar del coeficiente de fiabilidad de los datos que se analizan, incluso cuando el enfoque de su investigación no sea psicométrico.

Estudio 4: Los hallazgos de este estudio indican que las medidas de resistencia del tronco obtenidas de cinco pruebas de campo (BTC, DEE, BS, SB-L, SB-R e Ito) presentan valores de fiabilidad entre sesiones deficientes en estudiantes de Educación Secundaria Obligatoria (12-16 años). Aunque todas las medidas de resistencia del tronco exhibieron buenas puntuaciones de fiabilidad relativa (ICC> 0,75), la precisión de la medición de cada prueba de campo fue pobre (CV_{TE}> 10%). Además, se observó un sesgo sistemático entre sesiones en las pruebas de campo de resistencia dinámica del tronco (BTC y DEE). Los resultados del MDC_{95} indican que los cambios superiores a 31,4%, 42,4%, 44,5%, 51,4%, 54,6% y 67,5% para las pruebas BTC, DEE, BS, SB-L, SB-R e Ito después de una intervención pueden considerarse reales o verdaderos con un nivel de certeza del 95%. Solo la prueba BTC presentó un MDC_{95} lo suficientemente bajo (31,4%) como para ser considerado aceptable en estudiantes de Secundaria, ya que estudios previos han observado mejoras de aproximadamente el 40% en la resistencia del tronco después de haber completado un programa de entrenamiento de 6 semanas. El desarrollo de sesiones de familiarización supervisadas antes de realizar las pruebas (centrar la atención en las pruebas dinámicas) y animar fuertemente a los adolescentes a hacer siempre su máximo esfuerzo en cada prueba, puede ayudar a mejorar los valores de fiabilidad encontrados en este estudio.

Estudio 5: El estudio actual proporciona un análisis transversal de cinco pruebas de campo de resistencia del tronco en una gran muestra de niños y adolescentes de ambos sexos. Los hallazgos proporcionan una comprensión más clara de las tendencias esperadas en el desarrollo de la resistencia muscular del tronco durante el crecimiento y la maduración de los estudiantes. Existe una clara tendencia hacia el aumento de la resistencia del tronco a medida que los participantes son mayores o más maduros. Por otro lado, no se encontraron diferencias de rendimiento relacionadas con el sexo según la edad cronológica, pero sí se encontraron diferencias relacionadas con el sexo según el estado de maduración en la mayoría de las pruebas de resistencia. Este hecho no apoya la estratificación de las puntuaciones de las pruebas de resistencia del tronco por edad cronológica y sexo utilizadas en las baterías internacionales de pruebas de condición física juvenil.

En conclusión, los niños y niñas pre-púberes mostraron un desempeño similar en todas las pruebas de campo, cuando entraron en el estirón puberal las diferencias fueron relevantes (BTC, Ito, DEE), donde los chicos mostraron valores más altos que las chicas en todas las pruebas. Después del PVC, las diferencias se mantuvieron para los flexores-laterales (prueba SB) y los extensores del tronco (prueba DEE), y se igualaron para las pruebas de flexores (pruebas BTC e Ito) y extensores del tronco (prueba BS).

Por lo tanto, es importante considerar la maduración biológica, más que la edad cronológica, para la prescripción global de ejercicio en jóvenes en edad escolar, y así optimizar la adaptación al entrenamiento y minimizar el riesgo de lesiones relacionadas con la actividad. Por último, sería interesante incluir estrategias de intervención preventivas para evitar el desarrollo desequilibrado de la resistencia muscular del tronco durante el crecimiento, particularmente durante el inicio de la pubertad.

5.2. Limitaciones de la tesis y futuras líneas de investigación.

La presente tesis doctoral presenta ciertas limitaciones, y aunque la mayoría de ellas han sido abordas en sus respectivos estudios, a continuación, se presentan algunas limitaciones que podrían ser el punto de partida para nuevos estudios y proyectos de investigación.

Línea 1. Evolución del dolor de espalda y del "Fitness Postural" según el estado madurativo.

Los resultados de la presente tesis muestran que el dolor de espalda está asociado al desarrollo madurativo de los jóvenes, pero en una muestra no representativa de la población infantil y juvenil por lo que la extrapolación de los resultados a la población general no sería adecuada. Además, no fue posible establecer relaciones causa-efecto debido a que se empleó un diseño transversal. Por lo tanto, es necesario profundizar en dicha asociación y posible relación causa-efecto a través de estudios longitudinales y poder determinar de forma adecuada en qué estado madurativo existe una mayor vulnerabilidad.

Del mismo modo, para conocer cómo cambia el "Fitness Postural" a lo largo de los años y su asociación con el dolor de espalda, son necesarios estudios longitudinales desde edades tempranas para ver cómo evoluciona el morfotipo sagital de la columna vertebral, el rango de movimiento y la resistencia muscular del tronco. Además de incluir la valoración de la calidad de movimiento y conseguir un protocolo "Fitness Postural" más completo.

Por otro lado, estos estudios longitudinales deberían iniciarse en edades tempranas con niños que aún no han sufrido ningún tipo de dolor de espalda, y así poder establecer relaciones causa-efecto de forma correcta.

Además, a partir de los datos recogidos en la presente tesis, se pretenden desarrollar modelos de predicción novedosos para identificar con precisión a los jóvenes con alto o bajo riesgo de sufrir dolor de espalda. Para ello, se analizarán y compararán las capacidades individuales y combinadas de varias medidas obtenidas a través del protocolo "Fitness Postural" para predecir prospectivamente el riesgo de dolor de espalda utilizando un análisis basado en Redes Bayesianas.

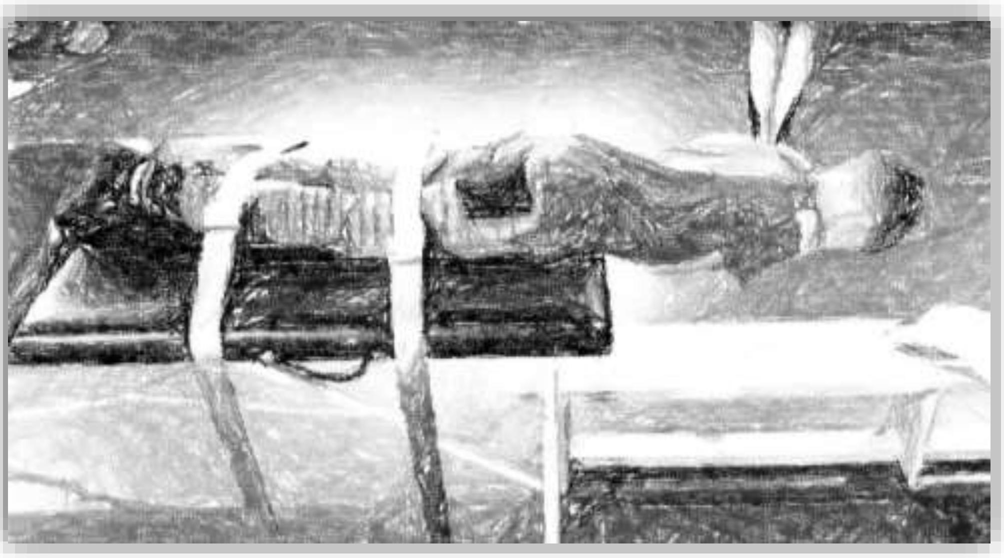
Línea 2. Validación y fiabilidad de las pruebas de resistencia de la musculatura del tronco en niños.

Los resultados de fiabilidad de las pruebas de campo para la resistencia de la musculatura del tronco no se han mostrado adecuados en términos de fiabilidad absoluta.

Por lo que sería necesaria la realización de nuevos estudios teniendo en cuenta las limitaciones del presentado en esta tesis. Nuevos estudios deben focalizar el objetivo en realizar y analizar los efectos de las sesiones de familiarización previas y así determinar el número de sesiones de familiarización necesarias para que los resultados obtenidos con estas pruebas sean correctos. Además, teniendo en cuenta que la fiabilidad es una propiedad que depende de las características de la muestra donde se aplique, es importante comprobar la fiabilidad de las pruebas utilizadas previamente con adolescentes, en niños más pequeños. En la literatura existente, son muy pocos los estudios que han investigado la fiabilidad de dichas pruebas en escolares de Educación Primaria, por ejemplo, por no decir que en niños más pequeños apenas existen estudios.

CHAPTER V

CONCLUSIONS



5.1. Conclusion

The main conclusions obtained from the studies that make up this doctoral thesis are presented below:

Research line 1:

Study 1: The present study showed a prevalence of BP associated with the maturational state and the weight of participants, finding different prevalence patterns after adjusted by sex. BP as well as LBP in boys, was associated with being overweight or obese, while in girls it was associated with greater pubertal development. The characteristics of BP were also associated with weight and maturation status, regardless of sex, finding a higher frequency and limitation of ADLs among older or overweight/obese participants. Therefore, the present study supports the statement that the prevalence of BP should be reported by spinal area, sex and stage of maturation and complemented with its characteristics, frequency and intensity as a minimum, to offer a complete view of the evolution and severity of BP in young population.

Weight is a modifiable risk factor whose optimization might generate benefits on the musculoskeletal health of schoolchildren and, although sex, age and state of maturity are biological and non-modifiable factors, they must all be taken into account at the time to carry out educational prevention programs.

Study 2: The present study showed that children and adolescents in pre- and circa-PHV presented sex-related differences in the assessment of "Postural Fitness" protocol. Boys presented a morphotype more "hyperkyphotic" along with a greater posterior pelvic tilt and girls a morphotype more "hyperlordotic" and more neutral pelvic tilt. On the other hand, boys had a hip ROM reduced and more endurance in the lateral-flexor muscles. Girls, however, showed more hip ROM and more extensor muscle endurance.

Concerning BP, associations were found between BP in boys and higher IMC, and between BP in girls and reduced ROM in hip flexor muscles and poor trunk flexor endurance. When analyzing the associations between "Postural Fitness" and LBP, boys with LBP showed an association with a higher BMI and a worse trunk lateral-flexors endurance, and girls with LBP showed association with a reduced ROM in hip flexor muscles and a worse endurance for trunk flexor and lateral-flexor muscles.

Hip ROM and trunk muscle endurance are modifiable risk factors whose optimization can generate benefits in the musculoskeletal health of children and adolescents, taking into account always their sex and biological maturity, as well as the alignment of the spine, at the time to carry out educational prevention programs.

Study 3: The main findings of the current RG meta-analysis report that only the trunk extension endurance measure obtained through the Biering-Sorensen test may present sufficient scientific evidence in terms of inter- (ICC+ = 0.94, based on 8 cohorts) and intra-tester (intra-session (ICC+ = 0.88, based on 12 cohorts) and inter-session (ICC+ = 0.88, based on 27 cohorts)) reliability to justify its use (mainly in male adults) for practical and clinical purposes. Furthermore, this trunk extension endurance measure may be even more reliable when a familiarization session with the Biering-Sorensen testing procedure is previously carried out and a plumb-line is used to control the participant's position during the test.

None of the rest of the trunk extension endurance measures obtained through the other static (prone isometric chest raise test and prone double straight-leg raise test) and dynamic field-based (dynamic extensor endurance test) tests showed sufficient scientific evidence (in terms of reliability) to promote their use for practical and clinical purposes. Therefore, more reliability studies conducted in different cohorts are needed before the use of these field-based tests can be recommended to assess trunk extensor endurance in both practical and research settings.

Finally, the high RI rate found in this study (72.8%) suggests that researchers should be more aware of the fact that RI is an erroneous practice that should be eradicated because it can cause errors in the estimation of the measures used. Thus, in future research, authors should report the reliability coefficient of the scores for the data being analyzed, even when the focus of their research is not psychometric.

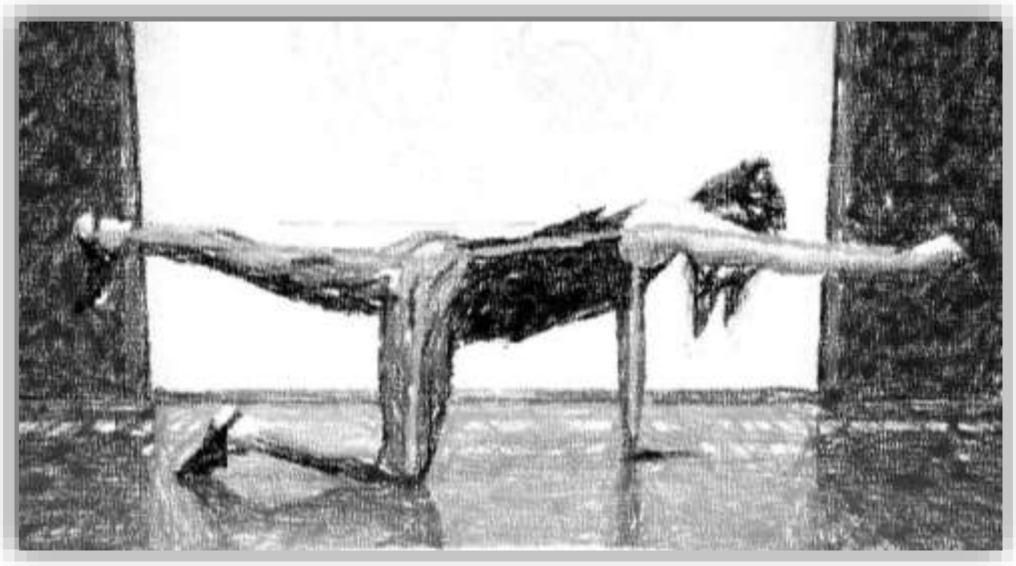
Study 4: To conclude, the findings from this study indicate that the trunk endurance measures obtained from five popular field-based tests (BTC, DEE, BS, SB-L, SB-R, and Ito) present poor inter-session reliability scores in high school-aged students (12-16 years). Even though all trunk endurance measures exhibited good relative reliability scores (ICC > 0.75), the precision of measurement of each field-based test was poor (CVTE > 10%). Furthermore, inter-session systematic bias was observed in the dynamic trunk endurance field-based tests (BTC and DEE). The MDC95 results indicate that changes higher than 31.4, 42.4, 44.5, 51.4, 54.6, and 67.5% for the BTC, DEE, BS, SB-L, SB-R, and Ito tests after an intervention may be considered as real or true with a 95% level of certainty. Only the BTC test obtained a MDC95 low enough (31.4%) to be considered as acceptable for high school-aged students as previous studies have reported improvements of approximately 40% in trunk endurance after having completed a 6-week training program. The use of supervised familiarization sessions before performing the tests (focus the attention on dynamic tests) and strongly encourage adolescents to do always their maximum effort in each test may help improve the reliability scores shown in this study.

Study 5: The current study provides a cross-sectional analysis of five trunk endurance field-based tests in a large sample of children and adolescents of both sexes. The findings may assist practitioners by providing a clearer understanding of expected trends in the development of trunk muscular endurance during student growth and maturation. There is a clear trend towards increasing trunk endurance as the participants are older or more mature. On the other hand, no sex-related performance differences were found according to chronological age, but sex-related differences were found according to maturational status in most of the endurance tests, this does not support the stratification of trunk endurance test scores by age used in international batteries of youth physical fitness tests.

In conclusion, pre-pubertal boys and girls showed similar performance in all field-based tests, when they entered the growth spurt the differences were relevant (BTC, Ito, DEE), with boys showing higher scores than girls in all tests. After the PHV, the differences were maintained for the lateral trunk flexors (SB test) and the DEE test and were equalized for the trunk flexor tests (BTC and Ito tests) and the BS test.

Therefore, as a result of the findings obtained, it is important to consider biological maturation, rather than chronological age, for the global exercise prescription for school-age youth, and thus optimize adaptation to training and minimize the risk of activity-related injuries. Finally, it would be interesting to include preventive intervention strategies to avoid the unbalanced development of trunk muscular endurance during growth, particularly during the onset of growth spurt.

REFERENCIAS BIBLIOGRÁFICAS



1. Hill, JJ.; Keating, JL. Risk factors for the first episode of low back pain in children are infrequently validated across samples and conditions: a systematic review. *J Physiother.* 2010, *56*(4), 237–44, doi: 10.1016/S1836-9553(10)70006-0.
2. Kamper, S.; Yamato, TP.; Williams, CM. The prevalence, risk factors, prognosis and treatment for back pain in children and adolescents: An overview of systematic reviews. *Best Pract Res Clin Rheumatol.* 2016, *30*(6), 1021–36, doi: 10.1016/j.berh.2017.04.003.
3. Calvo-Muñoz, I.; Gómez-Conesa, A.; Sánchez-Meca, J. Prevalence of low back pain in children and adolescents: A meta-analysis. *BMC Pediatr.* 2013, *13*(14), 1–12, doi: 10.1186/1471-2431-13-14.
4. Junge, T.; Wedderkopp, N.; Boyle, E.; Kjaer, P. The natural course of low back pain from childhood to young adulthood - a systematic review. *Chiropr Man Ther.* 2019, *27*(1), 1–10,
5. Jeffries, L.; Milanese, SF.; Grimmer-Somers, KA. Epidemiology of Adolescent Spinal Pain: A Systematic Overview of the Research Literature. *Spine (Phila Pa 1976).* 2007, *32*(23), 2630–7, doi: 10.1097/BRS.0b013e318158d70b.
6. Hoy, D.; Brooks, P.; Blyth, F.; Buchbinder, R. The Epidemiology of low back pain. *Best Pract Res Clin Rheumatol.* 2010, *24*(6), 769–81, doi: 10.1016/j.berh.2010.10.002.
7. Lardon, A.; Leboeuf-Yde, C.; Le Scanff, C.; Wedderkopp, N. Is puberty a risk factor for back pain in the young? A systematic critical literature review. *Chiropr Man Ther.* 2014, *22*(1), 1–12,
8. Beynon, AM.; Hebert, JJ.; Lebouef-Yde, C.; Walker, BF. Potential risk factors and triggers for back pain in children and young adults. A scoping review, part I: Incident and episodic back pain. *Chiropr Man Ther.* 2019, *27*(58), 1–7, doi: 10.1186/s12998-019-0280-9.
9. Hoy, D.; March, L.; Brooks, P.; Woolf, A.; Blyth, F.; Vos, T.; et al. Measuring the global burden of low back pain. *Best Pract Res Clin Rheumatol.* 2010, *24*(2), 155–65, doi: 10.1016/j.berh.2009.11.002.
10. Wirth, B.; Potthoff, T.; Rosser, S.; Humphreys, BK.; De Bruin, ED. Physical risk factors for adolescent neck and mid back pain: A systematic review. *Chiropr Man Ther.* 2018, *26*(1), 1–10,
11. Dionne, CE.; Dunn, KM.; Croft, PR.; Nachemson, AL.; Buchbinder, R.; Walker,

BF.; et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine (Phila Pa 1976)*. 2008, *33*(1), 95–103,

12. Lauridsen, HH.; Hestbaek, L. Development of the young spine questionnaire. *BMC Musculoskelet Disord*. 2013, *14*(185), 1–9,
13. Wirth, B.; Humphreys, BK. Pain characteristics of adolescent spinal pain. *BMC Pediatr*. 2015, *15*(42), 1–11,
14. Hestbaek, L.; Leboeuf-Yde, C.; Kyvik, KO.; Manniche, C. The course of low back pain from adolescence to adulthood: Eight-year follow-up of 9600 twins. *Spine (Phila Pa 1976)*. 2006, *31*(4), 468–72, doi: 10.1097/01.brs.0000199958.04073.d9.
15. Wirth, B.; Knecht, C.; Humphreys, K. Spine day 2012: Spinal pain in Swiss school children- epidemiology and risk factors. *BMC Pediatr*. 2013, *13*(159), 1–10, doi: 10.1186/1471-2431-13-159.
16. Aartun, E.; Hartvigsen, J.; Wedderkopp, N.; Hestbaek, L. Spinal pain in adolescents: prevalence, incidence, and course: a school-based two-year prospective cohort study in 1,300 Danes aged 11-13. *BMC Musculoskelet Disord*. 2014, *15*(187), 1–8,
17. Hasler, CC. Back pain during growth. *Swiss Med Wkly*. 2013, *143*(January),
18. Sato, T.; Ito, T.; Hirano, T.; Morita, O.; Kikuchi, R.; Endo, N.; et al. Low back pain in childhood and adolescence: a cross-sectional study in Niigata City. *Eur SPINE J*. 2008, *17*(11), 1441–7, doi: 10.1007/s00586-008-0788-5.
19. James, SL.; Abate, D.; Abate, KH.; Abay, SM.; Abbafati, C.; Abbasi, N.; et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018, *392*(10159), 1789–858,
20. Haas, M.; Goldberg, B.; Aickin, M.; Ganger, B.; Attwood, M. A practice-based study of patients with acute and chronic low back pain attending primary care and chiropractic physicians: Two-week to 48-month follow-up. *J Manipulative Physiol Ther*. 2004, *27*(3), 160–9,
21. Kovacs, FM.; Abraira, V.; Zamora, J.; Fernández, C. The Transition from Acute to Subacute and Chronic Low Back Pain. *Spine (Phila Pa 1976)*. 2005, *30*(15), 1786–92,

22. Maher, C.; Underwood, M.; Buchbinder, R. Non-specific low back pain. *Lancet*. 2017, *389*, 736–47,
23. Hartvigsen, J.; Hancock, MJ.; Kongsted, A.; Louw, Q.; Ferreira, ML.; Genevay, S.; et al. What low back pain is and why we need to pay attention. *Lancet*. 2018, *391*(10137), 2356–67,
24. Dunn, KM.; Hestbaek, L.; Cassidy, JD. Low back pain across the life course. *Best Pract Res Clin Rheumatol*. 2013, *27*(5), 591–600, doi: 10.1016/j.berh.2013.09.007.
25. Briggs, AM.; Smith, AJ.; Straker, LM.; Bragge, P. Thoracic spine pain in the general population: Prevalence, incidence and associated factors in children, adolescents and adults. A systematic review. *BMC Musculoskelet Disord*. 2009, *10*(1), 77, doi: 10.1186/1471-2474-10-77.
26. Scarabottolo, C.; Pinto, R.; Oliveira, C.; Zanuto, E.; Cardoso, J.; Christofaro, D. Back and neck pain prevalence and their association with physical inactivity domains in adolescents. *Eur Spine J*. 2017, *26*(9), 2274–80, doi: 10.1007/s00586-017-5144-1.
27. Kamper, S.; Henschke, N.; Hestbaek, L.; Dunn, K.; Williams, C. Musculoskeletal pain in children and adolescents. *Brazilian J Phys Ther*. 2016, *20*(3), 275–84,
28. Kjaer, P.; Wedderkopp, N.; Korsholm, L.; Leboeuf-Yde, C. Prevalence and tracking of back pain from childhood to adolescence. *BMC Musculoskelet Disord*. 2011, *12*(98), 1–11, doi: 10.1186/1471-2474-12-98.
29. Wedderkopp, N.; Andersen, LB.; Froberg, K.; Leboeuf-Yde, C. Back pain reporting pattern in a Danish population-based sample of children and adolescents. *BMC Musculoskelet Disord*. 2001, *6*(52), 1–5, doi: 10.1186/1471-2474-6-52.
30. Brattberg, G. Do pain problems in young school children persist into early adulthood? A 13-year follow-up. *Eur J Pain*. 2004, *8*(3), 187–99,
31. Jones, GT.; Silman, AJ.; Power, C.; Macfarlane, GJ. Are common symptoms in childhood associated with chronic widespread body pain in adulthood? Results from the 1958 British birth cohort study. *ARTHRITIS Rheum*. 2007, *56*(5), 1669–75,
32. Mirovsky, Y.; Jakim, I.; Halperin, N.; Lev, L. Non-specific back pain in children and adolescents: A prospective study until maturity. *J Pediatr Orthop Part B*. 2002, *11*(4), 275–8,

33. Jones, GT.; Watson, KD.; Silman, AJ.; Symmons, DPM.; Macfarlane, GJ. Predictors of low back pain in British schoolchildren: A population-based prospective cohort study. *Pediatrics*. 2003, *111*(4), 822–8,
34. Kovacs, F.; Gestoso, M.; Gil del Real, M.; Lopez, J.; Mufraggi, N.; Mendez, J. Risk factors for non-specific low back pain in schoolchildren and their parents: A population based study. *Pain*. 2003, *103*(3), 259–68, doi: 10.1016/S0304-3959(02)00454-2.
35. Potthoff, T.; De Bruin, ED.; Rosser, S.; Humphreys, BK.; Wirth, B. A systematic review on quantifiable physical risk factors for non-specific adolescent low back pain. *J Pediatr Rehabil Med*. 2018, *11*(2), 79–94,
36. Calvo-Muñoz, I.; Kovacs, FM.; Roqué, M.; Gago, I.; Seco, J. Risk Factors for Low Back Pain in Childhood and Adolescence: A Systematic Review [Internet]. Vol. 34, The Clinical Journal of Pain. F.M. Kovacs, Kovacs Back Pain Unit, HLA-University Hospital, Avda. Valladolid No. 83, Madrid, Spain, United States: Lippincott Williams and Wilkins; 2018 [cited 2020 May 12]. 468–484 p., doi: 10.1097/AJP.0000000000000558.
37. Beynon, AM.; Hebert, JJ.; Lebouef-Yde, C.; Walker, BF. Potential risk factors and triggers for back pain in children and young adults. A scoping review, part II: unclear or mixed types of back pain. *Chiropr Man Ther*. 2019, *27*(61), 1–12,
38. Ardakani, EM.; Leboeuf-Yde, C.; Walker, BF. Failure to define low back pain as a disease or an episode renders research on causality unsuitable: Results of a systematic review. *Chiropr Man Ther*. 2018, *26*(1), 1–10,
39. Bielza, C.; Larrañaga, P. Discrete Bayesian Network Classifiers: A Survey. *ACM Comput Surv*. 2014, *47*(1), 1–43,
40. Faigenbaum, AD.; Lloyd, RS.; Oliver, JL. Essentials of Youth Fitness. Champaign, USA: Human Kinetics; 2019.,
41. Poussa, MS.; Heliövaara, MM.; Seitsamo, JT.; Könönen, MH.; Hurmerinta, KA.; Nissinen, MJ. Anthropometric measurements and growth as predictors of low-back pain: A cohort study of children followed up from the age of 11 to 22 years. *Eur Spine J*. 2005, *14*(6), 595–8,
42. Hebert, JJ.; Leboeuf-Yde, C.; Franz, C.; Lardon, A.; Hestbæk, L.; Manson, N.; et al. Pubertal development and growth are prospectively associated with spinal pain in young people (CHAMPS study-DK). *Eur Spine J*. 2019, *28*(7), 1565–71,
43. Lloyd, RS.; Oliver, JL.; Faigenbaum, AD.; Myer, GD.; De Ste Croix, MB.

Chronological age vs. biological maturation: implications for exercise programming in youth. *J Strength Cond Res*. 2014, *28*(5), 1454–64,

44. Malina, RM.; Bouchard, C.; Bar-Or, O. Growth, Maturation, and Physical Activity. 2nd ed. Champaign, USA: Human Kinetics; 2004. 712 p.,
45. Shultz, SJ.; Nguyen, AD.; Schmitz, RJ. Differences in lower extremity anatomical and postural characteristics in males and females between maturation groups. *J Orthop Sports Phys Ther*. 2008, *38*(3), 137–49,
46. Leboeuf-Yde, C.; Kyvik, KO. At what age does low back pain become a common problem? A study of 29,424 individuals aged 12-41 years. *Spine (Phila Pa 1976)*. 1998, *23*(2), 228–34, doi: 10.1097/00007632-199801150-00015.
47. Janssens, KAM.; Rosmalen, JGM.; Ormel, J.; Verhulst, FC.; Hunfeld, JAM.; Mancl, LA.; et al. Pubertal status predicts back pain, overtiredness, and dizziness in american and dutch adolescents. *Pediatrics*. 2011, *128*(3), 553–9,
48. LeResche, L.; Mancl, LA.; Drangsholt, MT.; Saunders, K.; Von Korff, M. Relationship of pain and symptoms to pubertal development in adolescents. *Pain*. 2005, *118*(1–2), 201–9,
49. Wedderkopp, N.; Andersen, LB.; Froberg, K.; Leboeuf-Yde, C. Back pain reporting in young girls appears to be puberty-related. *BMC Musculoskelet Disord*. 2005, *6*, 52, doi: 10.1186/1471-2474-6-52.
50. Malina, RM.; Rogol, AD.; Cumming, SP.; Coelho E Silva, MJ.; Figueiredo, AJ. Biological maturation of youth athletes: Assessment and implications. *Br J Sports Med*. 2015, *49*(13), 852–9,
51. Mirwald, RL.; Baxter-Jones, ADG.; Bailey, DA.; Beunen, GP. An assessment of maturity from anthropometric measurements. *Med Sci Sports Exerc*. 2002, *34*(4), 689–94, doi: 10.1097/00005768-200204000-00020.
52. Moore, SA.; McKay, HA.; Macdonald, H.; Nettlefold, L.; Baxter-Jones, ADG.; Cameron, N.; et al. Enhancing a somatic maturity prediction model. *Med Sci Sports Exerc*. 2015, *47*(8), 1755–64,
53. Fransen, J.; Bush, S.; Woodcock, S.; Novak, A.; Deprez, D.; Baxter-Jones, ADG.; et al. Improving the Prediction of Maturity From Anthropometric Variables Using a Maturity Ratio. *Pediatr Exerc Sci*. 2018, *30*(2), 296–307,
54. Sjolie, AN. Persistence and change in nonspecific low back pain among adolescents: a 3-year prospective study. *Spine (Phila Pa 1976)*. 2004, *29*(21), 2452–7,

55. Jones, GT.; Macfarlane, GJ. Predicting Persistent Low Back Pain in Schoolchildren: A Prospective Cohort Study. *Arthritis Care Res (Hoboken).* 2009, *61*(10), 1359–66,
56. Mikkelsson, M.; El-Metwally, A.; Kautiainen, H.; Auvinen, A.; Macfarlane, GJ.; Salminen, JJ.; et al. Onset, prognosis and risk factors for widespread pain in schoolchildren: A prospective 4-year follow-up study. *Pain.* 2008, *138*(3), 681–7, doi: 10.1016/j.pain.2008.06.005.
57. Smith, A.; Beales, D.; Sullivan, PO.; Bear, N.; Straker, L. Low back pain with impact at 17 years of age is predicted by early adolescent risk factors from multiple domains: analysis of the Western Australian pregnancy cohort (raine) study. *J Orthop Sports Phys Ther.* 2017, *47*(10), 752–62, doi: 10.2519/jospt.2017.7464.
58. Dario, AB.; Kamper, SJ.; O'Keeffe, M.; Zadro, J.; Lee, H.; Wolfenden, L.; et al. Family history of pain and risk of musculoskeletal pain in children and adolescents: a systematic review and meta-analysis. *Pain.* 2019, *160*(11), 2430–9,
59. Hestbaek, L.; Iachine, IA.; Leboeuf-Yde, C.; Kyvik, KO.; Manniche, C.; L., H.; et al. Heredity of Low Back Pain in a Young Population: A Classical Twin Study. *Twin Res.* 2004, *7*(1), 16–26, doi: 10.1375/13690520460741408.
60. Batley, S.; Aartun, E.; Boyle, E.; Hartvigsen, J.; Stern, PJ.; Hestbæk, L. The association between psychological and social factors and spinal pain in adolescents. *Eur J Pediatr.* 2019, *178*(3), 275–86,
61. Stallknecht, S.; Strandberg-Larsen, K.; Hestbæk, L.; Andersen, A-MMN. Spinal pain and co-occurrence with stress and general well-being among young adolescents: a study within the Danish National Birth Cohort. *Eur J Pediatr.* 2017, *176*(6), 807–14, doi: 10.1007/s00431-017-2915-y.
62. Goncalves, TR.; Mediano, MFF.; Sichieri, R.; Cunha, DB.; T.R., G.; M.F.F., M.; et al. Is health-related quality of life decreased in adolescents with back pain? *Spine (Phila Pa 1976).* 2018, *43*(14), E822–9, doi: 10.1097/BRS.0000000000002520.
63. Lardon, A.; Leboeuf-Yde, C.; Le Scanff, C. Is back pain during childhood or adolescence associated with muscle strength, muscle endurance or aerobic capacity: Three systematic literature reviews with one meta-analysis. *Chiropr Man Ther.* 2015, *23*(1), 1–12, doi: 10.1186/s12998-015-0065-8.

64. Andersen, LB.; Wedderkopp, N.; Leboeuf-Yde, C.; Bo Andersen, L.; Wedderkopp, N.; Leboeuf-Yde, C.; et al. Association between back pain and physical fitness in adolescents. *Spine (Phila Pa 1976)*. 2006, *31*(15), 1740–4, doi: 10.1097/01.brs.0000224186.68017.e0.

65. Fritz, JM.; Clifford, SN. Low Back Pain in Adolescents: A Comparison of Clinical Outcomes in Sports Participants and Nonparticipants. *J Athl Train*. 2010, *45*(1), 61–6,

66. Haag, T-B.; Villefort, C.; Beckmann, C.; Mayer, HM.; Schneider, C. Prevalence of back pain in adolescent pupils according to their activity. *Sport Sci Health*. 2016, *12*(3), 383–7, doi: 10.1007/s11332-016-0302-0.

67. Heneweer, H.; Vanhees, L.; Picavet, HSJ. Physical activity and low back pain: A U-shaped relation? *Pain*. 2009, *143*(1–2), 21–5,

68. Szita, J.; Boja, S.; Szilagyi, A.; Somhegyi, A.; Varga, PP.; Lazary, A. Risk factors of non-specific spinal pain in childhood. *Eur Spine J*. 2018, *27*(5), 1119–26, doi: 10.1007/s00586-018-5516-1.

69. Martínez-Crespo, G.; Rodríguez-Piñero, M.; López-Salguero, AII.; Zarco-Periñan, MJJ.; Ibáñez-Campos, T.; Echevarría-Ruiz de Vargas, C.; et al. Dolor de espalda en adolescentes: prevalencia y factores asociados. *Rehabilitación*. 2009, *43*(2), 72–80, doi: 10.1016/S0048-7120(09)70773-X.

70. Auvinen, J.; Tammelin, T.; Taimela, S.; Zitting, P.; Karppinen, J. Associations of physical activity and inactivity with low back pain in adolescents. *Scand. J. Med. Sci. Sports*. 2008, *18*(2), 188–94, doi: 10.1111/j.1600-0838.2007.00672.x.

71. Sato, T.; Ito, T.; Hirano, T.; Morita, O.; Kikuchi, R.; Endo, N.; et al. Low back pain in childhood and adolescence: Assessment of sports activities. *Eur. Spine J*. 2011, *20*(1), 94–9, doi: 10.1007/s00586-010-1485-8.

72. Sanz-Mengibar, J.M.; Sainz de Baranda, P.; Santonja, F. Training intensity and sagittal curvature of the spine in male and female artistic gymnasts. *J. Sports Med. Phys. Fitness*. 2018, *58*(4), 465–71, doi: 10.23736/S0022-4707.17.06880-3.

73. Sainz de Baranda, P.; Cejudo, A.; Moreno-Alcaraz, V.J.; Martinez-Romero, M.T.; Aparicio-Sarmiento, A.; Santonja-Medina, F. Sagittal spinal morphotype assessment in 8 to 15 years old Inline Hockey players. *PeerJ*. 2020, *8*(e8229), 1–31, doi: 10.7717/peerj.8229.

74. Ginés-Díaz, A.; Martínez-Romero, M.T.; Cejudo, A.; Aparicio-Sarmiento, A.; Sainz de Baranda, P. Sagittal Spinal Morphotype Assessment in Dressage and

Show Jumping Riders. *J. Sport Rehabil.* 2019, *29*(5), 533–40, doi: 10.1123/jsr.2018-0247.

75. Sitthipornvorakul, E.; Janwantanakul, P.; Purepong, N.; Pensri, P.; Van Der Beek, A.J. The association between physical activity and neck and low back pain: A systematic review. *Eur. Spine J.* 2011, *20*(5), 677-689, doi: 10.1007/s00586-010-1630-4.
76. Aartun, E.; Hartvigsen, J.; Boyle, E.; Hestbaek, L. No associations between objectively measured physical activity and spinal pain in 11-15-year-old Danes. *Eur. J. Pain (United Kingdom).* 2016, *20*(3), 447–57, doi: 10.1002/ejp.746.
77. Gill, D.K.; Davis, M.C.; Smith, A.J.; Straker, L.M. Bidirectional relationships between cigarette use and spinal pain in adolescents accounting for psychosocial functioning. *Br. J. Health Psychol.* 2014, *19*(1), 113–31, doi: 10.1111/bjhp.12039.
78. Mikkonen, P.; Leino-Arjas, P.; Remes, J.; Zitting, P.; Taimela, S.; Karppinen, J. Is smoking a risk factor for low back pain in adolescents? A prospective cohort study. *Spine (Phila Pa 1976).* 2008, *33*(5), 527–32, doi: 10.1097/BRS.0b013e3181657d3c.
79. Feldman, D.; Shrier, I.; Rossignol, M.; Abenhaim, L.; Ehrmann-Feldman, D.; Rossignol, M.; et al. Risk factors for the development of low back pain in adolescence. *Am. J. Epidemiol.* 2001, *154*(1), 30–6, doi: 10.1093/aje/154.1.30.
80. Shiri, R.; Karppinen, J.; Leino-Arjas, P.; Solovieva, S.; Viikari-Juntura, E. The association between obesity and low back pain: A meta-analysis. *Am. J. Epidemiol.* 2010, *171*(2), 135–54, doi: 10.1093/aje/kwp356.
81. Auvinen, J.P.; Tammelin, T.H.; Taimela, S.P.; Zitting, P.J.; Jarvelin, M.-R.; Taanila, A.M.; et al. Is insufficient quantity and quality of sleep a risk factor for neck, shoulder and low back pain? A longitudinal study among adolescents. *Eur. Spine J.* 2010, *19*(4), 641–9, doi: 10.1007/s00586-009-1215-2.
82. Mustard, C.A.; Kalcevich, C.; Frank, J.W.; Boyle, M. Childhood and early adult predictors of risk of incident back pain: Ontario Child Health Study 2001 follow-up. *Am. J. Epidemiol.* 2005, *162*(8), 779–86, doi: 10.1093/aje/kwi271.
83. Mikkonen, P.; Laitinen, J.; Remes, J.; Tammelin, T.; Taimela, S.; Kaikkonen, K.; et al. Association between overweight and low back pain: A population-based prospective cohort study of adolescents. *Spine (Phila Pa 1976).* 2013, *38*(12), 1026–33, doi: 10.1097/BRS.0b013e3182843ac8.

84. Sano, A.; Hirano, T.; Watanabe, K.; Endo, N.; Ito, T.; Naohito, T. Body mass index is associated with low back pain in childhood and adolescence: a birth cohort study with a 6-year follow-up in Niigata City, Japan. *Eur. Spine J.* 2015, *24*(3), 474–81, doi: 10.1007/s00586-014-3685-0.
85. Paulis, W.D.; Silva, S.; Koes, B.W.; Van Middelkoop, M. Overweight and obesity are associated with musculoskeletal complaints as early as childhood: A systematic review. *Obes. Rev.* 2014, *15*(1), 52–67, doi: 10.1111/obr.12067.
86. Palmer, A.; Poveda, J.; Martinez-Laguna, D.; Reyes, C.; de Bont, J.; Silman, A.; et al. Childhood overweight and obesity and back pain risk: a cohort study of 466997 children. *BMJ Open.* 2020, *10*(e036023), doi: 10.1136/bmjopen-2019-036023.
87. Dolphens, M.; Cagnie, B.; Coorevits, P.; Vanderstraeten, G.; Cardon, G.; D'Hooge, R.; et al. Sagittal standing posture and its association with spinal pain: A school-based epidemiological study of 1196 flemish adolescents before age at peak height velocity. *Spine (Phila Pa 1976).* 2012, *37*(19), 1657–66, doi: 10.1097/BRS.0b013e3182408053.
88. Dolphens, M.; Cagnie, B.; Coorevits, P.; Vleeming, A.; Danneels, L. Classification system of the normal variation in sagittal standing plane alignment: A study among young adolescent boys. *Spine (Phila Pa 1976).* 2013, *38*(16), 1003–12, doi: 10.1097/BRS.0b013e318280cc4e.
89. Smith, A.; O'Sullivan, P.; Straker, L. Classification of sagittal thoraco-lumbo-pelvic alignment of the adolescent spine in standing and its relationship to low back pain. *Spine (Phila Pa 1976).* 2008, *33*(19), 2101–7, doi: 10.1097/BRS.0b013e31817ec3b0.
90. Dolphens, M.; Vansteelandt, S.; Cagnie, B.; Vleeming A.; Nijs, J.; Vanderstraeten, G.; et al. Multivariable modeling of factors associated with spinal pain in young adolescence. *Eur. Spine J.* 2016, *25*(9), 2809–21, doi: 10.1007/s00586-016-4629-7.
91. Cavalcante Filho, D.E.A.; de Lima Viana, C.N.; de Santana Cabral, M.P.S.; Cunha, F.V.M.; de Sousa Pacheco, F.; da Silva Brito, A.K.; et al. Low back pain in adolescents: A school screening. *J. Hum. Growth Dev.* 2014, *24*(3), 347–53, doi: 10.7322/jhdg.88975.
92. O'Sullivan, P.B.; Smith, A.J.; Beales, D.; Straker, L.M. Association of Biopsychosocial Factors With Degree of Slump in Sitting Posture and Self-

Report of Back Pain in Adolescents: A Cross-Sectional Study. *Phys. Ther.* 2011, *91*(4), 470–83, doi: 10.2522/ptj.20100160.

93. Astfalck, R.G.; O'Sullivan, P.B.; Straker, L.M.; Smith, A.J.; Burnett, A.; Caneiro, J.P.; et al. Sitting Postures and Trunk Muscle Activity in Adolescents With and Without Nonspecific Chronic Low Back Pain: An Analysis Based on SubclassificatioN. *Spine (Phila Pa 1976).* 2010, *35*(14), 1387–95, doi: 10.1016/j.math.2009.12.007.
94. Santonja, F.; Pastor, A.; Serna, L. Valoración radiográfica de las desalineaciones sagitales del raquis. *Selección.* 2000, *9*(4), 216–29.
95. Serna, L.; Santonja, F.; Pastor, A. Exploración clínica del plano sagital del raquis. *Ortop y Deport.* 1996, *5*(2), 88–2.
96. Santonja, F. Exploración clínica y radiográfica del raquis sagital: sus correlaciones [Tesis doctoral]. Universidad de Murcia; 1993.
97. Willner, S. Spinal Pantograph-A Non-Invasive Technique for Describing Kyphosis and Lordosis in the Thoraco-Lumbar Spine. *Acta Orthop Scand.* 1981, *52*(5), 525–9,
98. Debrunner, H.U. *Diagnóstico ortopédico.* 2ª Edición. Toray: Barcelona, 1976.
99. Widhe, T. Spine: posture, mobility and pain. A longitudinal study from childhood to adolescence. *Eur. Spine J.* 2001, *10*(2), 118–23, doi: 10.1007/s005860000230.
100. Oksanen, A.; Salminen, J.J. Tests of spinal mobility and muscle strength in the young: Reliability and normative values. *Physiother. Theory Pract.* 1996, *12*(3), 151–60, doi: 10.3109/09593989609036431.
101. Wojtys, E.M.; Ashton-Miller, J.A.; Huston, L.J.; Moga, P.J. The association between athletic training time and the sagittal curvature of the immature spine. *Am. J. Sports Med.* 2000, *28*(4), 490–8, doi: 10.1177/03635465000280040801.
102. Mannion, A.F.; Knecht, K.; Balaban, G.; Dvorak, J.; Grob, D. A new skin-surface device for measuring the curvature and global and segmental ranges of motion of the spine: Reliability of measurements and comparison with data reviewed from the literature. *Eur. Spine J.* 2004, *13*(2), 122–36, doi: 10.1007/s00586-003-0618-8.
103. Ayala, F.; Sainz De Baranda, P.; Cejudo, A.; Santonja, F. Pruebas angulares de estimación de la flexibilidad isquiosural: descripción de los procedimientos exploratorios y valores de referencia. *Rev. Andaluza Med. del Deport.* 2013, *6*(3), 120–8, doi: 10.1016/S1888-7546(13)70046-7.

104. Cejudo, A.; Sainz de Baranda, P.; Ayala, F.; De Ste Croix, M.; Santonja-Medina, F. Assessment of the Range of Movement of the Lower Limb in Sport: Advantages of the ROM-SPORT I Battery. *Int. J. Environ. Res. Public Health*. 2020, *17*(20), 7606, doi: 10.3390/ijerph17207606.

105. Kanchanomai, S.; Janwantanakul, P.; Pensri, P.; Jiamjarasrangsi, W. A Prospective Study of Incidence and Risk Factors for the Onset and Persistence of Low Back Pain in Thai University Students. *ASIA-PACIFIC J. PUBLIC Heal.* 2015, *27*(2), NP106–15, doi: 10.1177/1010539511427579.

106. Jones, M.A.A.; Stratton, G.; Reilly, T.; Unnithan, V.B.B. Biological risk indicators for recurrent non-specific low back pain in adolescents. *Br. J. Sports Med.* 2005, *39*(3), 137–40, doi: 10.1136/bjsm.2003.009951.

107. Zawadka, M.; Skublewska-Paszkowska, M.; Gawda, P.; Lukasik, E.; Smolka, J.; Jablonski, M. What factors can affect lumbopelvic flexion-extension motion in the sagittal plane?: A literature review. *Hum. Mov. Sci.* 2018, *58*(March), 205–18, doi: 10.1016/j.humov.2018.02.008.

108. Balague, F.; Bibbo, E.; Melot, C.; Szpalski, M.; Gunzburg, R.; Keller, T.S.; et al. The association between isoinertial trunk muscle performance and low back pain in male adolescents. *Eur. Spine J.* 2010, *19*(4), 624–32, doi: 10.1007/s00586-009-1168-5.

109. Aartun, E.; Hartvigsen, J.; Hestbaek, L. Validity of Commonly Used Clinical Tests to Diagnose and Screen for Spinal Pain in Adolescents: A School-Based Cohort Study in 1300 Danes Aged 11-15 Years. *J. Manipulative Physiol. Ther.* 2016, *39*(2), 76–87, doi: 10.1016/j.jmpt.2016.01.007.

110. Sjolie, A.N.; Ljunggren, A.E. The significance of high lumbar mobility and low lumbar strength for current and future low back pain in adolescents. *Spine (Phila Pa 1976).* 2001, *26*(23), 2629–36, doi: 10.1097/00007632-200112010-00019.

111. Faries, M.D.; Greenwood, M. Core Training: Stabilizing the Confusion. *Strength Cond. J.* 2007, *29*(2), 10–25, doi: 10.1519/00126548-200704000-00001.

112. McGill, S.M.; Childs, A.; Liebenson, C. Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. *Arch. Phys. Med. Rehabil.* 1999, *80*(8), 941–4, doi: 10.1016/S0003-9993(99)90087-4.

113. Abdelraouf, O.R.; Abdel-aziem, A.A. The relationship between core endurance and back dysfunction in collegiate male athletes with and without nonspecific

low back pain. *Int. J. Sports Phys. Ther.* 2016, *11*(3), 337–44.

114. Smith, J.J.; Eather, N.; Morgan, P.J.; Plotnikoff, R.C.; Faigenbaum, A.D.; Lubans, D.R. The Health Benefits of Muscular Fitness for Children and Adolescents: A Systematic Review and Meta-Analysis. *Sport Med.* 2014, *44*(9), 1209–23, doi: 10.1007/s40279-014-0196-4.

115. Geldhof, E.; Cardon, G.; De Bourdeaudhuij, I.; De Clercq, D. Back posture education in elementary schoolchildren: A 2-year follow-up study. *Eur. Spine J.* 2007, *16*(6), 841–50, doi: 10.1007/s00586-006-0227-4.

116. Perry, M.; Straker, L.M.; O'Sullivan, P.; Smith, A.; Hands, B. Fitness, Motor Competence, and Body Composition Are Weakly Associated With Adolescent Back Pain. *J. Orthop. Sport Phys. Ther.* 2009, *39*(6), 439–49, doi: 10.2519/jospt.2009.3011.

117. Salminen, J.J.; Maki, P.; Oksanen, A.; Pentti, J. Spinal mobility and trunk muscle strength in 15-year-old schoolchildren with and without low-back pain. *Spine (Phila Pa 1976).* 1992, *17*(4), 405–11, doi: 10.1097/00007632-199204000-00005.

118. Biering-Sorensen, F. Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine (Phila Pa 1976).* 1984, *9*(2), 106–19.

119. Ito, T.; Shirado, O.; Suzuki, H.; Takahashi, M.; Kaneda, K.; Strax, T.E. Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. *Arch. Phys. Med. Rehabil.* 1996, *77*, 75–9, doi: 10.1016/S0003-9993(96)90224-5.

120. McIntosh, G.; Wilson, L.; Affleck, M.; Hall, H. Trunk and lower extremity muscle endurance: Normative data. *J. Rehabil. Outcomes Meas.* 1998, *2*(4), 20–39.

121. Schellenberg, K.L.; Lang, J.M.; Chan, K.M.; Burnham, R.S. A clinical tool for office assessment of lumbar spine stabilization endurance: prone and supine bridge maneuvers. *Am. J. Phys. Med. Rehabil.* 2007, *86*(5), 380–6, doi: 10.1097/PHM.0b013e318032156a.

122. Luoto, S.; Heliövaara, M.; Hurri, H.; Alaranta, H. Static back endurance low-back pain and the risk of low-back pain. *Clin. Biomech.* 1995, *10*(6), 323–4.

123. Bliss, L.S.; Teeple, P. Core stability: the centerpiece of any training program. *Curr. Sports Med. Rep.* 2005, *4*(3), 179–83, doi: 10.1007/s11932-005-0064-y.

124. Robertson, L.D.; Magnusdottir, H. Evaluation of criteria associated with

abdominal fitness testing. *Res. Q. Exerc. Sport*. 1987, *58*(4), 355–9, doi: 10.1080/02701367.1987.10608112.

125. Jetté, M.; Sidney, K.; Cicutti, N. A critical analysis of sit-ups: a case for the partial curl-up as a test of abdominal muscular endurance. *CAHPER J*. 1984, *51*(1), 4–9.

126. Knudson, D.; Johnston, D. Validity and Reliability of a Bench Trunk-Curl Test of Abdominal Endurance. *J. Strength Cond. Res*. 1995, *9*(3), 165–9.

127. Brotons-Gil, E.; García-Vaquero, M.; Peco-González, N.; Vera-Garcia, F. Flexion-Rotation Trunk Test To Assess Abdominal Muscle Endurance: Reliability, Learning Effect, and Sex Differences. *J. Strength Cond. Res*. 2013, *27*(6), 1602–8, doi: 10.1519/JSC.0b013e31827124d9.

128. Merati, G.; Negrini, S.; Carabalona, R.; Margonato, V.; Veicsteinas, A. Trunk muscular strength in pre-pubertal children with and without back pain. *Pediatr. Rehabil*. 2004, *7*(2), 97–103, doi: 10.1080/13638490310001654754.

129. Mcgill, S.M. Low Back Stability: From Formal Description to Issues for Performance and Rehabilitation. *Exerc. Sport Sci. Rev*. 2001, *29*(1), 26–31, doi: 10.1097/00003677-200101000-00006.

130. Sadler, S.G.; Spink, M.J.; Ho, A.; De Jonge, X.J.; Chuter, V.H. Restriction in lateral bending range of motion, lumbar lordosis, and hamstring flexibility predicts the development of low back pain: A systematic review of prospective cohort studies. *BMC Musculoskelet. Disord*. 2017, *18*(1), 1–15, doi: 10.1186/s12891-017-1534-0.

131. Bernard, J-C.C.; Bard, R.; Pujol, A.; Combey, A.; Boussard, D.; Begue, C.; et al. Muscle assessment in healthy teenagers. Comparison with teenagers with low back pain. *Ann. Readapt. Med. Phys*. 2008, *51*(4), 274–83, doi: 10.1016/j.annrmp.2008.03.010.

132. Astfalck, R.G.; O’Sullivan, P.B.; Straker, L.M.; Smith, A.J. A detailed characterisation of pain, disability, physical and psychological features of a small group of adolescents with non-specific chronic low back pain. *Man. Ther*. 2010, *15*(3), 240–7, doi: 10.1016/j.math.2009.12.007.

133. Calvo-Muñoz, I.; Kovacs, F.M.; Roqué, M.; Seco-Calvo, J. The association between the weight of schoolbags and low back pain among schoolchildren: A systematic review, meta-analysis and individual patient data meta-analysis. *Eur. J. Pain (United Kingdom)*. 2020, *24*(1), 91–109, doi: 10.1002/ejp.1471.

134. Alghadir, A.H.; Gabr, S.A.; Al-Eisa, E.S. Mechanical factors and vitamin D deficiency in schoolchildren with low back pain: Biochemical and cross-sectional survey analysis. *J. Pain Res.* 2017, *10*, 855–65, doi: 10.2147/JPR.S124859.

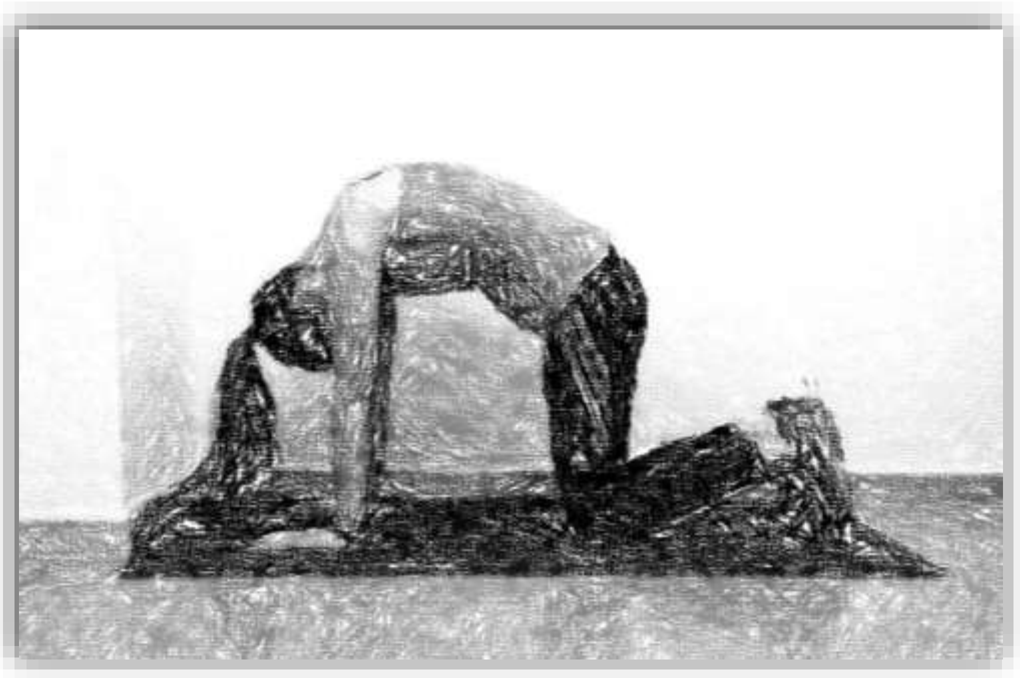
135. Erne, C.; Elfering, A. Low back pain at school: unique risk deriving from unsatisfactory grade in maths and school-type recommendation. *Eur. Spine J.* 2011, *20*(12), 2126–33, doi: 10.1007/s00586-011-1803-9.

136. Spiteri, K.; Busuttil, M.L.; Aquilina, S.; Gauci, D.; Camilleri, E.; Grech, V. Schoolbags and back pain in children between 8 and 13 years: a national study. *Br. J. Pain.* 2017, *11*(2), 81–6, doi: 10.1177/2049463717695144.

137. Delele, M.; Janakiraman, B.; Abebe, A.B.; Tafese, A.; van de Water, A.T.M. Musculoskeletal pain and associated factors among Ethiopian elementary school children. *BMC Musculoskelet. Disord.* 2018, *19*, doi: 10.1186/s12891-018-2192-6.

138. Dockrell, S.; Simms, C.; Blake, C. Schoolbag carriage and schoolbag-related musculoskeletal discomfort among primary school children. *Appl. Ergon.* 2015, *51*, 281–90, doi: 10.1016/j.apergo.2015.05.009.

ANEXOS



Appendix 1. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.

Section/Topic	#	Checklist Item	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	2–3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., web address), and, if available, provide registration information including registration number.	3 (PROSPERO ID: CRD42019123179)
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	3
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	3
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Table S2
Study selection	9	State the process for selecting studies (i.e., screening; eligibility; included in systematic review; and, if applicable, included in the meta-analysis).	3
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4 and Box S1
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4-5

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5 and Table S3
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4–5
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	6
Section/Topic	**#**	**Checklist Item**	**Reported on Page #**
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	Not applicable
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	6
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	6 Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7 and Table S4
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	6–9, Table 1 and Table S5
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group, and (b) effect estimates and confidence intervals, ideally with a forest plot.	10 Figure 2
Synthesis of results	21	Present results of each meta-analysis performed, including confidence intervals and measures of consistency.	Table 2
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15).	Not applicable
Additional analysis	23	Give results of additional analyses, if performed (e.g., sensitivity or subgroup analyses, meta-regression (see item 16)).	11–14 Tables 3, 4, and 5
DISCUSSION			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policymakers).	15–17

Limitations	25	Discuss limitations at the study and outcome level (e.g., risk of bias), and at the review-level (e.g., incomplete retrieval of identified research, reporting bias).	17
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, as well as implications for future research.	17
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	18

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097.

For more information, visit: www.prisma-statement.org.

Appendix 2. Search strategy (all databases).

Search Details [1]:	
S1	"Trunk muscle" OR "core muscle" OR "back muscle" OR "trunk extens*".
S2	"Endurance" OR "strength" OR "performance".
S3	S1 AND S2 AND test.
S4	S3 AND reliability.

[1] Limits: English and Spanish.

Appendix 3. Coding protocol.

General Study Descriptors

- Authors.
- Publication year.
- Language.
- Tester background (sports sciences, physical therapy, or medicine).
- Conflict of interest (yes or no).
- Continent (Europe, Africa, Asia, America, Oceania, Antarctica).
- Study objective (psychometric or not psychometric).
- Reliability analysis done with the same sample (yes or no).

Description of the Study Population

- Initial sample size.
- Final sample size.
- Sample type (children and adolescents or adults).
- Age and standard deviation.
- Gender distribution (% female).
- Gender (males, females, or males and females).
- Target population (asymptomatic or clinical).
- Physical activity level (sedentary, recreationally active, or athlete).
- Percentage attrition.

Description of the Field-Based Tests

- Test (Biering-Sorensen test, prone isometric chest raise test, prone double straight-leg raise test, or dynamic extensor endurance test).
- Test version (original or modified).
- Validated modification (yes or no).
- Tool (test bench or roman chair).
- Hand position (crossed on the chest, along the body, at the level of the ears).
- Part of the body on the edge (upper border of the iliac crest, anterior superior iliac spine (ASIS), not reported).
- Test duration (exhaustion or stopped in a certain time).
- Cadence (if it is applicable).
- Position control systems (visually, inclinometer, stadiometer, light sensor, plumb-line).
- Test score from the total sample (mean and SD).
- Test score from the reliability sample (mean and SD).

Type of Reliability Analysis

- Reliability (inter-tester or intra-tester (intra-session or internal consistency or inter-session or stability)).
- Intraclass correlation coefficient.

Characteristics of the Study Design

- Familiarization session (yes or no).
- Number of measurements.
- Number of testers.
- Time interval between measurements.
- Test conditions (type of administration, environment, instructions, etc. (yes or no)).

Appendix 4. Description of the consensus-based standards for the selection of health measurement instruments (COSMIN) methodology

Steps	Description
1. COSMIN risk of bias check-list	The checklist contains standards referring to design requirements and preferred statistical methods of studies on measurement properties. For each measurement property, a COSMIN box was developed containing all standards needed to assess the quality of a study on that specific measurement property. Each standard of the box is rated as "very good", "adequate", "doubtful", or "inadequate" quality. The overall rating of the quality of each study is determined by taking the lowest rating of any standard in the box.
2. Criteria for good measurement properties	The results of each study on a measurement property should be rated against the criteria for good measurement properties. Each result is rated as either sufficient (+ = ICC ≥ 0.70), insufficient (− = ICC < 0.70), or indeterminate (? = ICC not reported).
3. Summarize the evidence and grade the quality of the evidence	The results from different studies on one measurement property can be quantitatively pooled in a meta-analysis or qualitatively summarized. After pooling or summarizing all evidence per measurement property, and rating the pooled or summarized results against to the criteria for good measurement properties, the quality of the evidence is graded (high, moderate, low, very low evidence) on the basis of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. The GRADE approach uses five factors to determine the quality of the evidence: risk of bias (i.e., the methodological quality of the studies), inconsistency (i.e., unexplained inconsistency of results across studies), indirectness (i.e., evidence from different populations, interventions, or outcomes than the ones of interest), imprecision (i.e., total sample size of the available studies), and publication bias (i.e., negative results are less often published). The fifth factor, publication bias, is difficult to assess in studies on measurement properties, because of a lack of registries for this type of studies. Therefore, we do not take this factor into account in this meta-analysis [40].

Appendix 5. Individual characteristics of each study selected

a) Inter-Tester Reliability

Biering-Sorensen Test						
First Author, Year, and Country	Participant Characteristics	Test Description	Measurement Protocol	Reliability Coefficient	Conclusion [a]	Mean Results
Arab et al. (2007) [18] Iran	30 asymptomatic subjects (15 females) Mean age: 40.5 years	Modification of the original version. From the upper border of iliac crest and arms across the chest. Until exhaustion. Visual position control.	Two testers 15 min apart	$ICC_{2,1} = 0.78$	Poor	35.5 ± 7 s
Larsson et al. (2015) [24] Sweden	37 asymptomatic engineer soldiers (4 females) Mean age: 26 years	Modification of the original version. From the ASIS and hands at the level of the ears. Until exhaustion. Visual position control.	Four testers measured at the same time the same group of participants	$ICC_{2,1} = 0.99$	Very good	100.5 ± 29.2 s
Latimer et al. (1999) [23] Australia	(a) 23 adults with current low back pain (LBP) (13 females) Mean age: 35.9 years (b) 20 adults with previous LBP (9 females) Mean age: 36.6 years (c) 20 asymptomatic adults (10 females) Mean age: 28.8 years	Modification of the original version. From the ASIS and arms across the chest. Until exhaustion. Position control with inclinometer.	Two testers 15 min apart	a) $ICC_{1,1} = 0.88$ (0.73–0.95) b) $ICC_{1,1} = 0.77$ (0.52–0.9) c) $ICC_{1,1} = 0.83$ (0.62–0.93)	(a) Acceptable (b) Poor (c) Acceptable	(a) 94.6 ± 33.4 s (b) 107.7 ± 36.4 s (c) 132.6 ± 42.2 s

Palacín-Marín et al. (2013) [66] Spain	15 adults with LBP (9 females) Mean age: 37 years	Modification of the original version. From the upper border of iliac crest and arms across the chest. Until exhaustion. Position control with an inclinometer.	One tester examined in a real-time scenario and other tester examined the video recorded	$ICC_{2,1} = 0.92$ (0.91–0.93)	Very good	62.1 ± 32.6 s
Simmonds et al. (1998) [26] United States	(a) 22 adults with LBP (14 females) Mean age: 42.60 years (b) 48 asymptomatic adults (27 females) Mean age: 35.4 years	Modification of the original version. From the upper border of iliac crest and arms along the body. Until exhaustion. Visual position control.	Two testers measured twice within the test session	(a) $ICC_{1,1} = 0.99$ (b) $ICC_{1,1} = 0.99$	(a) Very good (b) Very good	(a) 45.9 ± 43 s (b) 77.8 ± 36.7 s
Prone Isometric Chest Raise Test						
Arab et al. (2007) [18] Iran	30 asymptomatic adults (15 females) Mean age: 40.5 years	Modification of the original version. Until exhaustion. Visual position control.	Two testers 15 min apart	$ICC_{2,1} = 0.9$	Very good	46 ± 13.5 s
Prone Double Straight-Leg Raise Test						
Arab et al. (2007) [18] Iran	30 asymptomatic adults (15 females) Mean age: 40.5 years	Original version. Until exhaustion. Visual position control.	Two testers 15 min apart	$ICC_{2,1} = 0.83$	Acceptable	36.5 ± 5.5 s

b) Intra-Tester and Intra-Session Reliability

Biering-Sorensen Test						
First Author, Year, and Country	Participant Characteristics	Test Description	Measurement Protocol	Reliability Coefficient	Conclusion [a]	Mean Results
Arab et al. (2007) [18] Iran	30 asymptomatic adults (15 females) Mean age: 40.5 years	Modification of the original version. From the upper border of iliac crest and arms across the chest. Until exhaustion. Visual position control.	Two measures 15 min apart	$ICC_{3,1}$ tester a = 0.8 $ICC_{3,1}$ tester b = 0.79	Tester a = acceptable Tester b = poor	35.5 ± 7 s
Demoulin et al. (2008) [43] Belgium	(a) 10 male university students Mean age: 22.3 years (b) 10 female university students Mean age: 20.7 years (c) 10 male asymptomatic adults Mean age: 41.1 years (d) 10 female asymptomatic adults Mean age: 42 years	Modification of the original version. From the ASIS and arms across the chest. Until exhaustion. Position control with a stadiometer.	Two measures within the same session	(a) ICC_{SEM} = 0.967 (b) ICC_{SEM} = 0.943 (c) ICC_{SEM} = 0.852 (d) ICC_{SEM} = 0.966 (e) ICC_{SEM} = 0.931 (f) ICC_{SEM} = 0.947	(a) Very good (b) Very good (c) Acceptable (d) Very good (e) Very good (f) Very good	(a) 113 ± 27.9 s (b) 141.8 ± 37.5 s (c) 130.8 ± 36.5 s (d) 122.5 ± 27.3 s (e) 93.2 ± 29 s (f) 114.8 ± 27.3 s

	(e) 10 male adults with LBP Mean age: 39.2 years (f) 10 female adults with LBP Mean age: 45.9 years					
Moffroid et al. (1993) [25] United States	28 female asymptomatic undergraduate students Mean age (no info available)	Modification of the original version. Arms along the body. Until exhaustion. Visual position control.	Two measures 15 min apart	$ICC_{NIA} = 0.87$	Acceptable	190.2 ± 66.3 s
Simmonds et al. (1998) [26] United States	(a) 44 adults with LBP (28 females) Mean age: 42.6 years (b) 48 asymptomatic adults (27 females) Mean age: 35.8 years	Not validated modification. From the upper border of iliac crest and arms along the body. Until exhaustion. Visual position control.	Two measures within the test session	(a) $ICC_{1,1} = 0.91$ (b) $ICC_{1,1} = 0.73$	(a) Desirable (b) Large	(a) 38.1 ± 34.9 s (b) 120.1 ± 75.3 s
Souza et al. (2016) [67] Brazil	48 female adults with LBP Mean age: 52 years	Validated modification. From the ASIS and arms across the chest. Until exhaustion. Position control with inclinometer.	Two measures 15 min apart	$ICC_{2,1} = 0.87$	Acceptable	54 ± 36 s

Arab et al. (2007) [18] Iran	30 asymptomatic adults (15 females) Mean age: 40.5 years	Validated modification. Until exhaustion. Visual position control.	Two measures 15 min apart	$ICC_{3,1}$ tester a = 0.9 $ICC_{3,1}$ tester b = 0.89	Tester a = desirable Tester b = acceptable	46 ± 13.5 s
Prone Double Straight-Leg Raise Test						
Arab et al. (2007) [18] Iran	30 asymptomatic adults (15 females) Mean age: 40.5 years	Original. Until exhaustion. Visual position control.	Two measures 15 min apart	$ICC_{3,1}$ tester a = 0.87 $ICC_{3,1}$ tester b = 0.85	Tester a = acceptable Tester b = acceptable	36.5 ± 5.5 s

c) Intra-Tester and Inter-Session Reliability

Biering-Sorensen Test						
First Author, Year, and Country	Participants Characteristics	Test Description	Measurement Protocol	Reliability Coefficient	Conclusion [a]	Mean Results
Dedering et al. (2000) [53] Sweden	10 asymptomatic adults (8 females) Mean age: 28.5 years	Modification of the original version. Hips flexed 40º and arms across the chest. Until exhaustion. Position control with sensor light.	Three measures 1 week apart	$ICC_{3,1}$ = 0.89	Acceptable	303.3 s
Dedering et al. (2010) [54] Sweden	15 adults with lumbar disc herniation (7 females) Mean age: 46 years	Modification of the original version. Hips flexed 40º and arms across the chest. Until exhaustion. Position control with sensor light.	Three measures 1 week apart	$ICC_{3,1}$ = 0.85	Acceptable	206 s

Demoulin et al. (2008) [43] Belgium	a) 10 male university students Mean age: 22.3 years b) 10 female university students Mean age: 20.7 years	Modification of the original version. From the ASIS and arms across the chest. Until exhaustion. Position control with a stadiometer.	Two measures 2 days apart	(a) ICC_{SEM} = 0.98 (b) ICC_{SEM} = 0.92	(a) Desirable (b) Desirable	a) 113 ± 27.9 s b) 141.8 ± 37.5 s
Demoulin et al. (2016) [56] Belgium	40 university students (20 females) Mean age: 21.1 years	Modification of the original version. Arms across the chest. Until exhaustion. Position control with stadiometer.	Two measures 3 days apart	$ICC_{3,1}$ = 0.87	Acceptable	157.2 ± 34.2 s
Geldhof et al. (2007) [57] Belgium	47 elementary schoolchildren Mean age: 10.1 years	Modification of the original version. From the upper border of the iliac crest and hands at the level of the ears. Until 240 seconds. Visual position control.	Two measures 1 week apart	$ICC_{3,1}$ = 0.63	Large	157.5 ± 53.75 s
Gruther et al. (2009) [8] Austria	21 adults with LBP Mean age: 43 years	Modification of the original version. From the upper border of the iliac crest and arms across the chest. Until exhaustion. Visual position control.	Two measures 2 weeks apart	$ICC_{3,1}$ = 0.59	Large	85.2 ± 49.3 s
Hannibal III et al. (2006) [27] United States	(a) 40 asymptomatic male adolescents Mean age: 15.1 years (b) 32 asymptomatic female adolescents Mean age: 15.5 years	Modification of the original version. Roman chair and arms across the chest. Until exhaustion.	Two measures Less than 1 week apart	(a) ICC_{ANOVA} = 0.99 (b) ICC_{ANOVA} = 0.99	(a) Desirable (b) Desirable	(a) 146.9 ± 65.1 s (b) 132.0 ± 46.0 s

		Position control with plumb-line.				
Jorgensen et al. (1986) [59] Denmark	10 asymptomatic male adults Mean age: 28 years	Modification of the original version. From the upper border of the iliac crest and arms across the chest. Until exhaustion. Visual position control.	Two measures 2 weeks apart	ICC_{NIA} = 0.89	Acceptable	277 ± 60.6 s
Juan-Recio et al. (2014) [60] Spain	27 asymptomatic male adults Mean age: 23.5 years	Modification of the original version. From the ASIS and arms across the chest. Until exhaustion. Visual position control.	Two measures 2 weeks apart	$ICC_{2,1}$ = 0.84 (0.72 - 0.91)	Acceptable	143.4 ± 42.5 s
Juan-Recio et al. (2018) [10] Spain	27 asymptomatic male adults Mean age: 24.1 years	Modification of the original version. From the ASIS and arms across the chest. Until exhaustion. Position control with stadiometer.	Two measures 1 month apart	$ICC_{2,1}$ = 0.78 (0.62 - 0.88)	Very large	136.92 ± 41.84 s
Keller et al. (2001) [61] Norway	(a) 31 adults with LBP (24 females) Mean age: 36 years (b) 31 asymptomatic adults (24 females) Mean age: 32 years	Modification of the original version. Roman chair, from pubis and arms across the chest. Until exhaustion. Visual position control.	Two measures 2 weeks apart	(a) $ICC_{1,k}$= 0.93 (b) $ICC_{1,k}$= 0.8	(a) Desirable (b) Acceptable	(a) 95.5 s (b) 138 s
Larsson et al. (2015) [24] Sweden	20 male ranger soldiers Mean age: 24 years	Modification of the original version. From the ASIS and hands at the level of the ears. Until exhaustion.	Two measures 1 week apart	$ICC_{3,1}$ = 0.85 (0.66–0.94)	Acceptable	100.5 ± 29.2 s

		Visual position control.				
--------- Lin et al. (2013) [62] Taiwan	10 asymptomatic male adolescents Mean age: 16.8 years	Modification of the original version. From the upper border of the iliac crest and arms across the chest. Until exhaustion. Visual position control.	Two measures No further information available	ICC_{NIA} = 0.75	Very large	93.0 ± 29.8 s
Mannion et al. (1997) [63] United Kingdom	10 asymptomatic adults Mean age: 26.10 years	Modification of the original version. From the upper border of the iliac crest and hands at the level of the ears. Until exhaustion. Visual position control.	Two measures 1 day apart	ICC_{ANOVA} = 0.98	Desirable	192.4 ± 76.7 s
Mayer et al. (1995) [9] United States	12 asymptomatic male adults Mean age: 30.20 years	Modification of the original version. Roman chair, from the pubis and arms across the chest. Until exhaustion. Visual position control.	Two measures 2 days apart	ICC_{NIA} = 0.2	Poor	No info available
Moffroid et al. (1994) [64] United States	(a) 32 adults with LBP (18 females) Mean age: 29.3 years (b) 9 Active adults with LBP (4 females) Mean age: 30.5 years (c) 20 inactive adults with LBP (13 females) Mean age: 28.1 years	Modification of the original version. Arms across the chest. Until exhaustion. Position control with inclinometer.	Two measures 1 day apart	(a) ICC_{ANOVA} = 0.82 (b) ICC_{ANOVA} = 0.96 (c) ICC_{ANOVA} = 0.39	(a) Acceptable (b) Desirable (c) Moderate	a) 48.6 ± 34.01 s b) 69.3 ± 59.5 s c) 40.6 ± 14.9 s
Ozcan Kahraman et al. (2016) [65] Turkey	38 adults with LBP (14 females) Mean age: 35 years	Modification of the original version. Arms across the chest.	Two measures 2 days apart	$ICC_{2,1}$ = 0.88 (0.79–0.93)	Acceptable	32.2 ± 24.6 s

		Until exhaustion. Visual position control.				
Palacín-Marín et al. (2013) [66] Spain	15 adults with LBP (9 females) Mean age: 37 years	Modification of the original version. From the upper border of iliac crest and arms across the chest. Until exhaustion. Position control with an inclinometer.	The same tester examined the video recording at 1 month after the first assessment	$ICC_{2,1}$ = 0.94 (0.93–0.95)	Desirable	62.1 ± 32.6 s
Simmonds et al. (1998) [26] United States	(a) 44 adults with LBP (28 females) Mean age: 42.60 years (b) 48 asymptomatic adults (27 females) Mean age: 35.4 years	Modification of the original version. From the upper border of iliac crest and arms along the body. Until exhaustion. Visual position control.	Two measures 2 weeks apart	(a) $ICC_{1,1}$ = 0.88 (b) $ICC_{1,1}$ = 0.68	(a) Acceptable (b) Very large	a) 38.1 ± 34.9 s b) 120.1 ± 75.3 s
Teyhen et al. (2011) [68] United States	64 asymptomatic adults (11 females) Mean age: 25.2 years	Modification of the original version. From the ASIS and arms across the chest. Until 240 seconds. Position control with inclinometer.	Two measures 2 days apart	$ICC_{2,1}$ = 0.79 (0.67–0.87)	Very large	108.6 ± 39.6 s
Waldhelm and Li (2012) [69] United States	15 asymptomatic male undergraduate students Mean age: 21.20 years	Modification of the original version. Arms across the chest. Until exhaustion. Visual position control.	Two measures 1 week apart	$ICC_{2,1}$ = 0.79 (0.38–0.93)	Very large	83.3 ± 29.4 s
Dynamic Extensor Endurance Test						
Hannibal III et al., (2006) [27] United States	(a) 40 asymptomatic male adolescents Mean age: 15.1 years (b) 32 asymptomatic female adolescents	(1) Modification of the original version. Roman chair, ROM of 90º and arms across the chest. Until exhaustion, 20 rep/min.	Two measures Less than 1 week apart	(1a) ICC_{ANOVA} = 0.99 (1b) ICC_{ANOVA} = 0.99	(1a) Desirable (1b) Desirable (2a) Desirable (2b) Desirable	(1a) 57.1 ± 34.3 reps (1b) 50.6 ± 22.6 reps

	Mean age: 15.5 years	Position control with plumb-line. (2) Modification of the original version. ROM of 90º and arms across the chest. Until exhaustion, 20 rep/min. Visual position control.		(2a) ICC_{ANOVA} = 0.99 (2b) ICC_{ANOVA} = 0.99		(2a) 53.4 ± 29.9 reps (2b) 44.0 ± 19.8 reps
Lin et al. (2013) [62] Taiwan	10 asymptomatic male adolescents Mean age: 16.8 years	Modification of the original version. ROM of 30º and arms across the chest. Until exhaustion. Visual position control.	Two measures 1 week apart	ICC_{NIA} = 0.92	Desirable	27.0 ± 10.3 reps
Prone Isometric Chest Raise Test						
del Pozo-Cruz et al. (2014) [55] Spain	(a) 12 male adults with LBP Mean age: 45.8 years (b) 19 female adults with LBP Mean age: 46 years	Modification of the original version. Until exhaustion. Visual position control.	Two measures 1 week apart	(a) ICC_{NIA} = 0.97 (b) ICC_{NIA} = 0.96	(a) Desirable (b) Desirable	(a) 82.4 ± 25.2 s (b) 84.3 ± 31.1 s
Ito et al. (1996) [11] Japan	(a) 90 asymptomatic adults (53 females) Mean age: 45.5 years (b) 100 adults with LBP (60 females) Mean age: 45.4 years	Original version. Until 300 seconds. Visual position control.	Two measures 3 days apart	(a) ICC_{NIA} = 0.97 (b) ICC_{NIA} = 0.93	(a) Desirable (b) Desirable	(a) 160.3 ± 70.5 s (b) 76.3 ± 53.9 s
Javadian et al. (2012) [58] Iran	15 adults with LBP Range age 18 - 45 years	Original version. Until 300 seconds. Visual position control.	Two measures 2 days apart	ICC_{NIA} = 0.85	Acceptable	22.5 ± 4.5 s

Biering-Sorensen test (Biering-Sorensen, 1984) is evaluated by measuring how many seconds the participant was able to keep the unsupported upper part of the body (from the upper border of the iliac crest) horizontal while placed prone with the buttocks and legs fixed to the table bench by three wide canvas straps, and the arms across the chest. The test is continued until the participant could no longer control his/her posture for a maximum of 240s. Prone isometric chest raise test (Ito et al., 1996) is assessed by measuring how many seconds the participant was able to keep the sternum off the floor while placed lying in a prone position, with the arms along the body. A small pillow is placed under the iliac crest to decrease the lumbar lordosis. The subject is asked to maintain the positions for as long as possible, not exceeding a 5 minute time limit. Prone double straight-leg raise test (McIntosh, Wilson, Affleck, and Hall, 1998) is evaluated with the participant in a prone position with hips extended, the hands underneath the forehead, and the arms perpendicular to the body. The participant is instructed to raise both legs until knee clearance is achieved. The test is continued until the participant could no longer able to maintain knee clearance. Dynamic extensor endurance test (Luoto, Heliövaara, Hurri, and Alaranta, 1995) is performed with the subject in a prone position with the unsupported upper part of the body (from the upper border of the iliac crest). The arms are positioned along the body and the buttocks and legs are fixed by three straps. With the spine kept straight, the subject is instructed to extend the trunk to neutral and then to lower the upper body 45 degrees. A repeated beat guided the subject to maintain a cadence of 25 repetitions per minute until exhaustion. The number of repetitions accomplished by the subject is counted. [a] ICC values were interpreted according to the following criteria: <0.1, trivial; 0.1–0.29, small; 0.3–0.49, moderate; 0.5–0.69, large; 0.7–0.89, very large; 0.9–1, nearly perfect. (a–f) indicate different cohorts in the same study, ASIS: anterior superior iliac spine, ROM: range of motion, $ICC_{2,1}$: intraclass correlation coefficient based on two-way random effects model, $ICC_{1,1}$: intraclass correlation coefficient based on one-way random effects model, $ICC_{1,k}$: intraclass correlation coefficient based on one-way random effects model, $ICC_{3,1}$: intraclass correlation coefficient based on two-way mixed effects model, ICC_{NIA}: no information available about intraclass correlation coefficient calculation, ICC_{ANOVA}: intraclass correlation coefficient calculated from results obtained from analysis of variance (ANOVA), ICC_{SEM}: intraclass correlation coefficient calculated from standard error of measurement scores and using standardized equations (see method section for more information),s: seconds, reps: repetitions.

Appendix 6. COSMIN risk of bias checklist (Box 6 reliability) and criteria for good measurement properties.

d) Inter-Tester Reliability

Study	Standards for Assessing the Risk of Bias					Final Score	Good Measurement Properties
	1	2	3	4	8		
Biering-Sorensen Test							
Arab et al., 2007 [18]	VG	D	A	VG	VG	D	+
Larsson et al., 2015 [24]	VG	VG	VG	VG	VG	VG	+
Latimer et al., 1999 [23]	A	D	A	VG	D	D	+
Palacín-Marín et al., 2013 [66]	VG	VG	VG	VG	VG	VG	+
Simmonds et al., 1998 [26]	A	D	A	VG	VG	D	+
Prone Isometric Chest Raise Test							
Arab et al., 2007 [18]	VG	D	A	VG	VG	D	+
Prone Double Straight-Leg Raise Test							
Arab et al., 2007 [18]	VG	D	A	VG	VG	D	+

e) Intra-Tester Reliability and Intra-Session Reliability

Study	Standards for Assessing the Risk of Bias					Final Score	Good Measurement Properties
	1	2	3	4	8		
Biering-Sorensen Test							
Arab et al., 2007 [18]	VG	D	A	VG	VG	D	+
Demoulin et al., 2008 [43]	A	D	A	VG	D	D	+
Moffroid et al., 1993 [25]	A	D	A	A	D	D	+
Simmonds et al., 1998 [26]	A	D	A	VG	VG	D	+
Souza et al., 2016 [67]	A	D	A	A	D	D	+
Prone Isometric Chest Raise Test							
Arab et al., 2007 [18]	VG	D	A	VG	VG	D	+
Prone Double Straight-Leg Raise Test							
Arab et al., 2007 [18]	VG	D	A	VG	VG	D	+

f) Intra-Tester Reliability and Inter-Session Reliability

Study	Standards for Assessing the Risk of Bias					Final Score	Good Measurement Properties
	1	2	3	4	8		
Biering-Sorensen Test							
Dedering et al., 2000 [53]	VG	D	VG	VG	VG	D	+
Dedering et al., 2010 [54]	VG	D	VG	VG	VG	D	+
Demoulin et al., 2008 [43]	A	D	A	VG	D	D	+
Demoulin et al., 2016 [56]	VG	D	VG	VG	VG	D	+
Geldhof et al., 2007 [57]	VG	D	VG	VG	VG	D	−
Gruther et al., 2009 [8]	VG	VG	VG	VG	VG	VG	−
Hannibal III et al., 2006 [27]	A	D	A	VG	D	D	+
Jorgensen & Nicolaisen 1986 [59]	VG	VG	VG	VG	VG	VG	+
Juan-Recio et al., 2014 [60]	VG	D	VG	VG	VG	D	+

Juan-Recio et al., 2018 [10]	VG	VG	VG	VG	VG	VG	+
Keller et al., 2001 [61]	VG	VG	VG	VG	VG	VG	+
Larsson et al., 2015 [24]	VG	D	VG	VG	VG	D	+
Lin et al., 2013 [62]	A	D	A	A	I	I	+
Mannion et al., 1997 [63]	A	D	A	VG	I	I	+
Mayer et al., 1995 [9]	A	D	A	A	D	D	−
Moffroid et al., 1994 [64]	VG	D	VG	VG	D	D	−
Ozkan Kahraman et al., 2016 [65]	VG	D	VG	VG	D	D	+
Palacín-Marín et al., 2013 [66]	VG	VG	VG	VG	VG	VG	+
Simmonds et al., 1998 [26]	A	VG	A	VG	VG	A	+
Teyhen et al., 2011 [68]	A	D	A	VG	D	D	+
Waldhelm & Li 2012 [69]	A	D	A	VG	D	D	+
Dynamic Extensor Endurance Test							
Hannibal III et al., 2006 [27]	A	D	A	VG	D	D	+
Lin et al., 2013 [62]	A	D	A	A	I	I	+
Prone Isometric Chest Raise Test							
del Pozo-Cruz et al., 2014 [55]	A	D	A	D	D	D	+
Ito et al., 1996 [11]	A	D	A	A	D	D	+
Javadian et al., 2012 [58]	A	D	A	A	D	D	+

Standards for assessing the risk of bias:

1) Patients were stable in the interim period on the construct to be measured (very good (VG): evidence provided that patients were stable; adequate (A): assumable that patients were stable; doubtful (D): unclear if patients were stable; inadequate (I): patients were not stable).

2) The time interval was appropriate (VG: time interval appropriate; D: doubtful whether time interval was appropriate or time interval was not stated; I: time interval not appropriate).

3) The test conditions were similar for the measurements (e.g., type of administration, environment, instructions) (VG: test conditions were similar (evidence provided); A: assumable that test conditions were similar; D: unclear if test conditions were similar; I: test conditions were not similar).

4) For continuous scores: an intraclass correlation coefficient (ICC) was calculated (VG: ICC calculated and model or formula of the ICC is described; A: ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change occurred; D: Pearson or Spearman correlation coefficient calculated without evidence provided that no systematic change occurred or with evidence that systematic change occurred; I: no ICC or Pearson or Spearman correlations calculated).

5) There were any other important flaws in the design or statistical methods of the study (VG: no other important methodological flaws; D: other minor methodological flaws; I: other important methodological flaws).

Criteria for good measurement properties (reliability): (+) ICC ≥ 0.70; (-) ICC < 0.70; (?) ICC not reported