The relationship between thoracic kyphosis and age, and normative values across age groups: a systematic review of healthy adults

Mattia Zappalà 1,3*, Stephen Lightbourne 2 and Nicola R. Heneghan 3

Abstract

Background: Thoracic kyphosis is reported to increase with ageing. However, this relationship has not been systematically investigated. People’s kyphosis often exceeds 40°, but 40° is the widely accepted cut-off and threshold for normality. Consequently, patients may be misclassified. Accurate restoration of kyphosis is important to avoid complications following spinal surgery. Therefore, specific reference values are needed. The objective of the review is to explore the relationship between thoracic kyphosis and age, provide normative values of kyphosis for different age groups and investigate the influence of gender and ethnicity.

Methods: Two reviewers independently conducted a literature search, including seven databases and the Spine Journal, from inception to April 2020. Quantitative observational studies on healthy adults (18 years of age or older) with no known pathologies, and measuring kyphosis with Cobb’s method, a flexicurve, or a kyphometer, were included. Study selection, data extraction, and study quality assessment (AQUA tool) were performed independently by two reviewers. The authors were contacted if clarifications were necessary. Correlation analysis and inferential statistics were performed (Microsoft Excel). The results are presented narratively. A modified GRADE was used for evidence quality assessment.

Results: Thirty-four studies (24 moderate-quality, 10 high-quality) were included (n = 7633). A positive moderate correlation between kyphosis and age was found (Spearman 0.52, p < 0.05, T5-T12). People’s kyphosis resulted greater than 40° in 65% of the cases, and it was significantly smaller in individuals younger than 40 years old (x < 40) than in those older than 60 years old (x > 60) 75% of the time (p < 0.05). No differences between genders were found, although a greater kyphosis angle was observed in North Americans and Europeans.

Conclusion: Kyphosis increases with ageing, varying significantly between x < 40 and x > 60. Furthermore, kyphosis appears to be influenced by ethnicity, but not gender. People’s thoracic sagittal curvature frequently exceeds 40°.

* Correspondence: mattia.zappala.89@gmail.com

1Physiotherapy Department, St John & St Elizabeth Hospital, 60 Grove End Rd, St John’s Wood, London, UK
2Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, College of Life and Environmental Sciences, University of Birmingham, Birmingham, UK
3Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, College of Life and Environmental Sciences, University of Birmingham, Birmingham, UK

Full list of author information is available at the end of the article

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Trial registration: The review protocol was devised following the PRISMA-P Guidelines, and it was registered on PROSPERO (CRD42020175058) before study commencement.

Keywords: Kyphosis, Hyperkyphosis, Ageing, Normative value, Correlation, Thoracic spine, Gender, Ethnic group, Reference values, Healthy adults

Background
Kyphosis, the convex curvature of the thoracic spine is considered ‘normal’ between 20 and 40° [1]. Where this exceeds 40°, the curvature is described as hyperkyphosis. This is associated with a higher risk of falling, developing pulmonary dysfunctions, and poor quality of life [2–4]. Hyperkyphosis is also associated with a higher risk of mortality for any cause [2–4]. A prospective longitudinal study, which followed 610 women for over 13 years, found that people with a greater thoracic kyphosis, who previously sustained a vertebral fracture, have a 1.5 times higher risk of death than those who have a smaller kyphotic curvature [5]. Consequently, it has been suggested that thoracic kyphosis is an important parameter to monitor, especially in the elderly population, to detect more frail people who may be at higher risk of unfavourable health [5].

The prevalence of hyperkyphosis increases with age; 20–40% of people older than 60 years of age and 55% of those older than 70 years have a kyphosis exceeding 40° [2–4]. Consequently, hyperkyphosis has been associated with ageing [4]. However, the relationship between kyphosis and age has not been systematically investigated. Individual studies show conflicting results [6, 7], and evidence supporting this association is derived from narrative reviews [2–4], rather than methodologically rigorous systematic reviews [8].

Despite evidence suggesting that people’s kyphosis often exceeds 40° [2–4], this value is widely used in clinical practice as the cut-off for normality [3, 4]. Consequently, clinicians may find many of their patients present with hyperkyphosis. Several authors have highlighted the need for a more accurate threshold for diagnosing hyperkyphosis [2–4], and a recent narrative review proposed to move the cut-off of normality to 50° [2]. The Scoliosis Research Society suggests using a range of 20–60° instead [9]. However, since people of different age groups have different degrees of kyphosis [2, 3], moving the cut-off of normality to a higher value, or expanding its range, may not reduce the risk of misdiagnosis. For these reasons, and due to the importance of the thoracic curvature when restoring patients’ sagittal alignment during spinal corrective surgery, to avoid post-operative complications such as proximal junctional kyphosis [10], having specific age-related reference values of kyphosis may be useful.

Objective
This systematic review aims to investigate the sagittal curvature of the thoracic spine of adults with no health conditions which may affect their thoracic kyphosis and do the following:

1. Explore the relationship between kyphosis and age
2. Provide reference values of kyphosis for different age groups
3. Examine data for differences between genders or ethnic groups

Methods
Protocol and registration
The review’s protocol followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for protocols (PRISMA-P) [11] and was registered on PROSPERO (CRD42020175058). The methods were informed by the Cochrane Handbook [12]. The manuscript adhered to the PRISMA [13] and the Synthesis Without Meta-analysis (SWiM) guidelines [14] for reporting.

Eligibility criteria
The research question was informed by the Sample, Phenomenon of Interest, Design, Evaluation, Research type (SPIDER) tool [15], whose details are in Table 1.

Information sources
Two reviewers (MZ/SL) independently searched for eligible articles on MEDLINE, EMBASE and PsycINFO through Ovid, and on AMED, The Index of Chiropractic Literature and CINAHL through EBESCO, from inception to April 2020. The Spine Journal, the reference list of the studies included in the review, and grey literature on SIGLE, through Open Grey, were also searched. The research was limited to studies published in English.

Search
Keyword selection was informed by scoping review and researcher expertise (NRH). The search strategy was individualised for each database, combining keywords, Medical Subject Headings, and Boolean operators, and following consultation with a librarian. Keywords selected were middle back, dorsal spine, middle spine, mid-back, thoracic spine, kyphosis, hyperkyphosis, Dowager’s hump, hunchback, rounded back, and sagittal curvature (see Additional file 1 for search strategy examples).
Table 1 Eligibility criteria

| Sample | Adults (18+ years old) without osteoporosis; vertebral fractures; pain; Scheuermann’s disease; scoliosis; history of spinal surgery or trauma; history of prolonged steroid use; rheumatological conditions; cardiac, lung or autoimmune diseases; cancer; metastasis; inflammatory or neurological disorders; pregnancy; or any genetic conditions affecting their bones, muscles or cartilage |
| Phenomenon of interest | Individuals’ thoracic sagittal alignment |
| Design | Any research design |
| Evaluation | Cobb’s method, a flexicurve or a Debrunner’s kyphometer |
| Research type | Quantitative, not performing interventions |

Study selection
The screening process was conducted independently by MZ and SL, then agreement was sought. In case of disagreement, a third reviewer (NRH) acted as a moderator. The studies were screened from their title and abstract first, then from their full text [8].

Data collection
The data collection process was informed by the Cochrane Handbook [16]. The data extraction form was piloted with data extraction performed independently by MZ and SL and then cross-checked. If further information was necessary to reach a consensus among the research team, the authors were contacted by MZ.

Data items
Data extraction was informed by the recommendations for reviews in clinical anatomy [8]. This included study title, author’s name, publication year, method for measuring kyphosis, degrees of kyphosis and range, sample size, age, age range, gender, body mass index, the standard deviation (SD) of the measures and ethnicity, defined as a group of people sharing cultural, geographical and social attributes.

Risk of bias in individual studies
The studies’ quality assessment was performed independently by MZ and SL; NRH acted as a moderator in case of disagreement. The Anatomical Quality Assessment (AQUA) tool, devised for assessing the quality of anatomical studies [17], was used. As suggested by Chhapola et al. [18], a supplementary table to improve the tool’s performance was created (see Additional file 2). The AQUA tool is composed of 5 domains (i.e. objective(s) and subject characteristics, study design, methodology characterisation, descriptive anatomy, reporting of results); each of them has a specific set of questions whose answers could be either yes, no or unclear to enable the readers to evaluate the study’s quality. Currently, only indications about how to evaluate each individual domain of the AQUA tool exist. To be considered at low risk of bias in a single domain, the study must receive yes answers to all the questions of that specific domain; otherwise, the study would be considered at high risk [17]. Each domain was evaluated following the procedure just described. However, since no guidance exists on how to classify the overall quality of the evaluated study, the research team agreed that for a study to be considered, overall, high-quality, this must be at low risk of bias in all five domains. If at low risk in three or four domains they were considered moderate-quality, otherwise low-quality. The tool was then piloted before study commencement by MZ and SL on five articles and inter-rater agreement computed according to McHugh [19]. Perfect agreement was achieved (κ = 1).

Summary measures
Data was analysed with Microsoft Excel of the Microsoft Office 365 package. Since kyphosis varies depending on the body references used to calculate it [6, 20], analysis was performed comparing the measurements for the same body references. The mean kyphosis and age were used for correlation analysis. Either the Pearson’s or Spearman’s correlation coefficient was computed, depending on whether the data were normally distributed or not. Data distribution was investigated with the Kolmogorov-Smirnov test, and correlation was interpreted as recommended [21]. The means and their precision estimates were used to calculate the reference/normative values, or ranges, of kyphosis for each age group. Since SDs represent the dispersion of the values around their means, whereas confidence intervals are used to assess a treatment’s efficacy [22], SDs were deemed to be more appropriate to establish ranges. The mean kyphosis was utilised for group comparisons. Previous evidence regarding the relationship between kyphosis and age [2, 4, 6] was used to create the groups for analysis. These were people younger than 40 years old (x < 40), people between 40 and 60 (40 < x < 60), people older than 60 (x > 60), people younger than 50 (x < 50), and those older than 50 years old (x > 50). Inferential statistics was performed using the independent two-tailed t-test, for two group comparisons (x < 50, x > 50), or one-way ANOVA, for multiple group comparison (x < 40, 40 < x < 60, x > 60). Gender and ethnic group differences were investigated comparing each individual age group using the independent two-tailed t-test. Levene’s test was used to assess between groups’ equality variances. The selected alpha level was 0.05, and the Bonferroni correction was applied for post hoc analysis, after ANOVA, to reduce the chances of type I error [23, 24].
Synthesis of results and risk of bias across studies

Since important clinical and methodological heterogeneity were observed during the scoping review, meta-analysis was not performed [25]. Data were synthesised narratively, and descriptive statistics presented [26]. The overall level of evidence was evaluated using a modified Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system [27]. Whilst limited to observational studies, if the results were consistent (> 80% concordant results) [28], precise, and obtained predominantly from high-quality studies, the overall quality was upgraded from low to moderate. For correlation analysis, consistency was assessed by evaluating the direction of the correlation (positive or negative). For the reference values and for gender and ethnic group comparisons, statistical significance between groups’ means was used. Correlation analysis to be precise must be statistically significant, whereas for the normative values and for gender and ethnic group comparisons, the ranges of the groups with statistically significant different means must not overlap. Furthermore, their difference must be greater than the standard error of measurements for the modality employed to calculate kyphosis. These values were 2.4° for the kyphometer [29], 0.4 cm for the flexicurve [30], and 3° for Cobb’s method [7]. If the results were inconsistent, imprecise and coming primarily from low-quality studies, the results’ quality was downgraded to very low.

Results

Study selection

A total of 12,366 studies were retrieved, and 68 selected for full-text screening. Thirty-eight studies were excluded after the full-text screening, and four added following reference review, resulting in a total of 34 studies included in the review [6, 7, 20, 31–61] (Fig. 1).

Study characteristics and individual studies results

Details about the included studies are in Table 2. From 7633 participants, the age range was 18–95 years old. Kyphosis was measured between C7–T12 (n = 220), T1–T12 (n = 2154), T2–T12 (n = 212), T3–T12 (n = 101), T4–T12 (n = 1617) and T5–T12 (n = 4018). Kyphosis was measured with a flexicurve in 293 individuals. Most studies used Cobb angle with just two (n = 293) studies using a flexicurve [47, 52].

Risk of bias within studies

Ten of the studies were high-quality [20, 31, 32, 36, 44–47, 57, 61], 24 were moderate-quality [6, 7, 33–35, 37–43, 48–56, 58–60] and none low-quality (see Table 3 for details). The most frequent limitation regarded studies’ methodology, with 12 studies [33, 34, 37–39, 42, 43, 48, 52, 53, 59] not reporting the accuracy of their measures. This limitation equally affected all measurement types.

Relationship between kyphosis and age

Only studies measuring kyphosis using Cobb’s method were included in the analysis because of the greater sample size, which provides greater statistical power [23], and those using a flexicurve included only women, limiting their generalisability. No analysis was performed for C7–T12 and T3–T12 because data came from single studies.

A positive correlation between kyphosis and age was found (see Table 4). The strength of the correlation was moderate for T5–T12 (Spearman 0.52) and low for T4–T12 (Spearman 0.45). The sample size for T5–T12 was more than double that for T4–T12 [25], giving more confidence in the findings for T5–T12.

Normative values

Table 4 provides details of the mean kyphosis and normative values of kyphosis for different age groups, as well as between-group mean difference in kyphosis and the sample sizes. The same studies utilised to investigate the relationship between kyphosis and age were also used for calculating the reference values. Only 12 studies divided their sample by age groups [6, 7, 20, 31, 32, 35, 41, 43, 48, 50, 58, 59]. The ranges surpassed 40° in people < 60 years old 58.3% of the time and 75% in those older, questioning the accuracy of the current cut-off for normality.

Gender and ethnic group differences

Fourteen studies specified sample ethnicity [20, 32, 34–38, 41–46, 59]; consequently, geographical provenience was the main determinant for ethnic group subdivision. Two studies were excluded from the sub-analysis between ethnicities. One study [60] did not divide their sample by age groups and did not report mean’s SD, whereas in the other study [48], the sample size was too small to exclude the chance of committing type II error. Fifteen of the included studies presented their results according to gender [6, 31, 32, 34, 36, 37, 40–43, 45, 48, 53, 58, 59], and only eight of those divided their sample by age [6, 31, 32, 41, 43, 48, 58, 59]. The results are reported in Table 4. No differences between genders were observed, but North Americans and Europeans showed a greater thoracic curvature than Asians (Fig. 2).

Synthesis of results

There is moderate-quality evidence that a moderate positive correlation between age and kyphosis exists and that kyphosis does not differ between genders. The quality of the evidence for the normative values presented,
and for the differences in kyphosis observed between ethnicities is low (Table 5).

Discussion
This is the first review exploring the relationship between kyphosis and age, in addition to providing normative kyphosis values for different ages, ethnic groups and genders. Findings evidence a positive correlation between kyphosis and age, as well as the influence of ethnicity on kyphosis. Gender, instead, does not appear to influence thoracic sagittal curvature.

Relationship between kyphosis and age
Muscle strength, vertebral body shape and intervertebral disc morphology can affect kyphosis angle [3]. However, vertebral body shape and intervertebral disc morphology account for 86–93% thoracic spine curvature [62]. Disc morphology has a stronger negative correlation with ageing than vertebral morphology [62, 63]. Therefore, the increase in thoracic kyphosis observed with ageing may be related to the changes occurring in intervertebral discs. Most of these changes occur in the middle section of the thoracic spine [64], which can explain why statistical significance was reached only when kyphosis was measured from T4/5. For these reasons, and due to the
| Authors and year      | Type of measurement | Results | Sample size | Age (SD) | Age range | Gender | Ethnicity | BMI (SD) | Study design                        |
|----------------------|---------------------|---------|-------------|----------|-----------|--------|-----------|----------|-------------------------------------|
| Amabile et al. 2016  | Cobb T1–T12         | 49 (±1.3) | 69          | 26.3     | 18–40     | Mix    | Europeans | 25.1 (±3.3) | Retrospective, observational        |
|                      | Cobb T4–T12         | 35.1 (±11.5) |             |          |           |        |           |          |                                     |
| Bakouny et al. 2017  | Cobb T1–T12         | 51 (±7.9) | 48          | 21.5     | 18–28     | Male   | Asians    | 21.8 (±2.4) | Prospective cross-sectional cohort study, observational |
|                      |                     | 47.7 (±9.1) | 44          | 21.5     | 18–28     | Female |          | 23.5 (±3.4) |                                     |
|                      |                     | 49.4 (±8.6) | 92          | 21.5     | 18–28     | Mix    |           |          |                                     |
|                      | Cobb T2–T12         | 50.8 (±8.8) |             |          |           |        |           |          |                                     |
|                      |                     | 47.4 (±10.1) |             |          |           |        |           |          |                                     |
|                      |                     | 49.2 (±9.5) |             |          |           |        |           |          |                                     |
|                      | Cobb T4–T12         | 43.1 (±8.8) |             |          |           |        |           |          |                                     |
|                      |                     | 41 (±10.4) |             |          |           |        |           |          |                                     |
|                      |                     | 42.1 (±9.6) |             |          |           |        |           |          |                                     |
| Bassani et al. 2019  | Cobb T1–T12         | 49.4 (±13.6) | 44          | 66       | 60–69     | Mix    | Europeans | 25 (±3)   | Prospective cross-sectional cohort study, observational |
|                      |                     | 54.4 (±13.2) | 83          | 74       | 70–79     |        |           |          |                                     |
|                      |                     | 63.3 (±11.3) | 27          | 83       | x > 79    |        |           |          |                                     |
|                      | Cobb T4–T12         | 40.8 (±10)  |             |          |           |        |           |          |                                     |
|                      |                     | 48 (±12.1)   |             |          |           |        |           |          |                                     |
|                      |                     | 57.7 (±10.8) |             |          |           |        |           |          |                                     |
| Endo et al. 2014     | Cobb T4–T12         | 27.5 (±9.6) | 86          | 35.9     | 23–59     | Mix    | Asians    | 21 (±2.7) | Prospective cross-sectional cohort study, observational |
| Endo et al. 2016     | Cobb T4–T12         | 30.5 (±8.3) | 30          | 35.2     | 22–50     | Male   | Asians    | 21 (±2.7) | Prospective cross-sectional cohort study, observational |
|                      |                     | 24.1 (±10.4) | 22          | 35.8     | 22–50     | Female |          | 21 (±2.7) |                                     |
|                      |                     | 27.8 (±9.7)  | 52          | 35.4     | 22–50     | Mix    |           | 21 (±2.7) |                                     |
| Gangnet et al. 2006  | Cobb T1–T12         | 44.5 (±11.3) | 34          | 30       |          | Mix    | Europeans | 25 (±3)   | Retrospective, observational        |
|                      | Cobb T4–T12         | 35.7 (±7.7)  |             |          |           |        |           |          |                                     |
| Gelb et al. 1995     | Cobb T5–T12         | 36 (±11)    | 27          | 40–49    |          | Mix    | North Americans | 20.9 (±1.49) | Prospective cross-sectional cohort study, observational |
|                      |                     | 32 (±10)    | 27          | 50–59    |          |        |           |          |                                     |
|                      |                     | 36 (±11)    | 37          | 60–69    |          |        |           |          |                                     |
|                      |                     | 33 (±14)    | 9           |          | x > 70   |        |           |          |                                     |
| Ganino et al. 2014   | Cobb T1–T12         | 29.9 (±5.12)| 10          | 24.6     | 22.27    | Female | South Americans | 26.2 (±2.32) | Prospective cross-sectional cohort study, observational |
|                      |                     | 36.11 (±3.71) | 10          | 43.5     | 22.88    |        |           |          |                                     |
|                      |                     | 37.14 (±2.67) | 10 | 62.4     | 2.67    |        |           |          |                                     |
| Hammerberg et al. 2003| Cobb T1–T12       | 52.5 (±12.2) | 50          | 76.3     | 70–85    | Mix    | North Americans | 21.1 (±2.4) | Prospective cross-sectional cohort study, observational |
| Hasagawa et al. 2016 | Cobb T1–T12         | 41.5 (±9.9) | 126         | 39.4     | 1.13     | Mix    | Asians    | 21.4 (±2.4) | Prospective cross-sectional cohort study, observational |
|                      | Cobb T4–T12         | 29.6 (±9.2)  |             |          |           |        |           |          |                                     |
| Hasagawa et al. 2017 | Cobb T1–T12         | 43.7 (±9)   | 40          | 40       |          | Male   | Asians    | 21.4     | Prospective cross-sectional cohort study, observational |
|                      |                     | 41 (±10.2)  | 96          | 39.6     |          | Female |          | 21.4     |                                     |
|                      |                     | 41.8 (±19.58) | 136 | 39.7     |          | Mix    |           | 21.4     |                                     |
| Hinman 2004          | Flexicurve          | 12.19 (±3.71) cm | 25          | 29.2     | 21–51    | Female | North Americans | 21.4     | Prospective cross-sectional cohort study, observational |
|                      |                     | 10.02 (±2.43) cm | 26          | 72.3     | 66–68    |        |           |          |                                     |
| Authors and year | Type of measurement | Results Kyphosis (SD) (degrees/cm) | Sample size | Age (SD) | Age range | Gender | Ethnicity | BMI (SD) | Study design |
|-----------------|---------------------|-----------------------------------|-------------|----------|-----------|--------|-----------|----------|--------------|
| Hu et al. 2016 [42] | Cobb T5–T12 | 24.6 (±9.4) | 161 | 23.2 (±3.4) | 18–45 | Male | Asians | | Prospective cross-sectional cohort study, observational |
| Hu et al. 2020 [33] | Cobb T5–T12 | 24.7 (±9.5) | 40 | 20–29 | Male | Asians | | | Prospective cross-sectional cohort study, observational |
| | | 23.8 (±8.5) | 46 | 20–29 | Female | | | | |
| | | 22.6 (±7.7) | 40 | 30–39 | Male | | | | |
| | | 23.9 (±8.5) | 41 | 30–39 | Female | | | | |
| | | 25.2 (±7.3) | 40 | 40–49 | Male | | | | |
| | | 27.2 (±8.5) | 42 | 40–49 | Female | | | | |
| | | 26.8 (±8.5) | 41 | 50–59 | Male | | | | |
| | | 27.4 (±8.5) | 45 | 50–59 | Female | | | | |
| | | 28.5 (±7.2) | 40 | 60–69 | Male | | | | |
| | | 27.5 (±8.5) | 44 | 60–69 | Female | | | | |
| | | 28.6 (±8.8) | 41 | 70–79 | Male | | | | |
| | | 30.8 (±13.2) | 40 | x > 80 | Male | | | | |
| | | 29 (±8.5) | 44 | x > 80 | Female | | | | |
| Iyer et al. 2016 [6] | Cobb T2–T12 | 46.3 (±6.3) | 5 | 21–30 | Male | North Americans | 26.5 (±6.6) | | Prospective cross-sectional cohort study, observational |
| | | 39.6 (±14.7) | 16 | 21–30 | Female | | | | |
| | | 41.2 (±13.3) | 21 | 21–30 | Mix | | | | |
| | | 37.3 (±10.7) | 6 | 31–40 | Male | | | | |
| | | 46 (±11) | 13 | 31–40 | Female | | | | |
| | | 43.2 (±11.4) | 19 | 31–40 | Mix | | | | |
| | | 59.6 (±21) | 4 | 41–50 | Male | | | | |
| | | 40.4 (±6.8) | 15 | 41–50 | Female | | | | |
| | | 44.5 (±13.2) | 19 | 41–50 | Mix | | | | |
| | | 48.5 (±13.4) | 3 | 51–60 | Male | | | | |
| | | 51.4 (±10.7) | 13 | 51–60 | Female | | | | |
| | | 50.9 (±11.2) | 16 | 51–60 | Mix | | | | |
| | | 52.4 (±11.3) | 11 | 61–70 | Male | | | | |
| | | 44.7 (±14.5) | 16 | 61–70 | Female | | | | |
| | | 47.7 (±13.6) | 27 | 61–70 | Mix | | | | |
| | | 46 (±15.7) | 9 | x > 71 | Male | | | | |
| | | 54.2 (±16.4) | 9 | x > 71 | Female | | | | |
| | | 49.8 (±16.1) | 18 | x > 71 | Mix | | | | |
| Iyer et al. 2016 [6] | Cobb T5–T12 | 34.8 (±7.7) | 5 | 21–30 | Male | North Americans | 26.5 (±6.6) | | Prospective cross-sectional cohort study, observational |
| | | 27 (±11.4) | 16 | 21–30 | Female | | | | |
| | | 28.9 (±11) | 21 | 21–30 | Mix | | | | |
| | | 27.8 (±9.9) | 6 | 31–40 | Male | | | | |
| | | 37.3 (±10.7) | 13 | 31–40 | Female | | | | |
| | | 31.2 (±9.3) | 19 | 31–40 | Mix | | | | |
| | | 46 (±22.3) | 4 | 41–50 | Male | | | | |
| | | 29.1 (±6.4) | 15 | 41–50 | Female | | | | |
| | | 32.7 (±12.8) | 19 | 41–50 | Mix | | | | |
| | | 39.9 (±12) | 3 | 51–60 | Male | | | | |
| | | 37.9 (±12.8) | 13 | 51–60 | Female | | | | |
| | | 38.3 (±12.2) | 16 | 51–60 | Mix | | | | |
| | | 39.5 (±8.7) | 11 | 61–70 | Male | | | | |
| | | 34.4 (±15.6) | 16 | 61–70 | Female | | | | |
| Authors and year       | Type of measurement | Results | Sample size | Age (SD) | Age range | Gender | Ethnicity | BMI (SD) | Study design                        |
|-----------------------|---------------------|---------|-------------|----------|-----------|--------|-----------|----------|-------------------------------------|
|                       | Kyphosis (SD)       |         |             |          |           |        |           |          |                                     |
|                       | (degrees/cm)        |         |             |          |           |        |           |          |                                     |
| Janssen et al. 2009   | Cobb T4–T12         | 36.4 ±1.3 | 27          | 61–70    | Mix       |        |           | 27.4 ±2.7 | Prospective cross-sectional cohort study, observational |
|                       |                     | 33.3 ±1.3 | 9           | > 71     | Male      |        |           | 26.8 ±4  |                                      |
|                       |                     | 44 ±1.6  | 9           | > 71     | Female    |        |           | 268 ±4   |                                      |
|                       |                     | 38.3 ±1.2 | 18          |          |           |        |           | 268 ±4   |                                      |
|                       |                     | 33.3 ±1.3 | 30          | 27       | Male      |        | Europeans | 219      |                                     |
|                       |                     | 44 ±1.6  | 184         | 21.2     | Male      |        | Asians    | 224 ±2.1 |                                     |
|                       |                     | 36.4 ±1.3 | 158         | 63.8     |           |        | Mix       | 239 ±2.9 |                                     |
|                       |                     | 44 ±1.6  | 99          | 52.7     |           |        | Mix       | 214      |                                     |
|                       |                     | 38.3 ±1.3 | 32.9       | 20–79    |           |        | Mix       | 28 ±6    | Retrospective, observational        |
|                       |                     | 36.4 ±1.3 | 27          |           |           |        |           |          |                                     |
|                       |                     | 33.3 ±1.3 | 139         | 50.8     |           |        |           |          |                                     |
|                       |                     | 44 ±1.6  | 26.8 ±4     |          |           |        |           |          |                                     |
|                       |                     | 38.3 ±1.3 | 70          |           |           |        |           |          |                                     |
| Pavlovic et al. 2013  | Flexicurve          | 3.2 ±2.2 (cm) | 104        | 23.4     | Male      |        | North Americans | 252 ±2.8 | Prospective cross-sectional cohort study, observational |
|                       |                     | 3.4 ±2.2 (cm) | 138        | 20–29    |           |        |           | 213 ±2.8 |                                     |
|                       |                     | 38.69 ±1.1 | 25          | 23.4     | Female    |        |           | 231 ±3.3 |                                     |
|                       |                     | 39 ±1.52  | 50          | 23.4     |           |        |           | 228 ±3.3 |                                     |
|                       |                     | 41 ±8.05  | 25          | 65.8     |           |        |           | 227 ±2.6 |                                     |
|                       |                     | 26.33 ±1.5 | 25          | 65.8     |           |        |           | 223 ±2.7 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 65.8     |           |        |           | 221 ±2.8 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 64.0     |           |        |           | 223 ±3.5 |                                     |
|                       |                     | 38.31 ±1.2 | 25          | 20–29    |           |        |           | 258 ±3.2 | Retrospective, observational        |
|                       |                     | 38.69 ±1.1 | 25          | 20–29    |           |        |           | 208 ±3.2 |                                     |
|                       |                     | 39 ±1.52  | 50          | 20–29    |           |        |           | 208 ±3.2 |                                     |
|                       |                     | 41 ±8.05  | 25          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 26.33 ±1.5 | 25          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 38.31 ±1.2 | 25          | 20–29    |           |        |           | 258 ±3.2 | Retrospective, observational        |
|                       |                     | 38.69 ±1.1 | 25          | 20–29    |           |        |           | 208 ±3.2 |                                     |
|                       |                     | 39 ±1.52  | 50          | 20–29    |           |        |           | 208 ±3.2 |                                     |
|                       |                     | 41 ±8.05  | 25          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 26.33 ±1.5 | 25          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 38.31 ±1.2 | 25          | 20–29    |           |        |           | 258 ±3.2 | Retrospective, observational        |
|                       |                     | 38.69 ±1.1 | 25          | 20–29    |           |        |           | 208 ±3.2 |                                     |
|                       |                     | 39 ±1.52  | 50          | 20–29    |           |        |           | 208 ±3.2 |                                     |
|                       |                     | 41 ±8.05  | 25          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 26.33 ±1.5 | 25          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 65.8     |           |        |           | 231 ±3.7 |                                     |
|                       |                     | 33.37 ±1.7 | 50          | 65.8     |           |        |           | 231 ±3.7 |                                     |
| Authors and year | Type of measurement | Results | Sample size | Age (SD) | Age range | Gender | Ethnicity | BMI (SD) | Study design |
|------------------|---------------------|---------|-------------|----------|-----------|--------|-----------|----------|--------------|
| Uehara et al. 2019 [41] | Cobb T5–T12 | 35 (±11.73) | 49 | 58.43 (±6.6) | 50–79 | Mix | | | |
| Urrutia et al. 2014 [60] | Cobb T5–T12 | 36 | 760 | 66.6 (±11) | Mix | South Americans | | | Prospective cross-sectional cohort study, observational |
| Vialle et al. 2005 [57] | Cobb T4–T12 | 40.6 (±10) | 300 | 35.4 (±12) | 20–70 | Mix | Europeans | 23.5 (±3) | Prospective cross-sectional cohort study, observational |
| Yang et al. 2017 [34] | Cobb T1–T12 | 36.72 (±11.11) | 183 | 50.52 | | | Male | Mix | Asians |
| Yukawa et al. 2018 [43] | Cobb T1–T12 | 34.8 (±10.3) | 34.8 (±10) | 34.8 (±10.3) | 34.8 (±10) | | | | |
| Zhu et al. 2014 [46] | Cobb T5–T12 | 34.8 (±10.5) | 34.8 (±10.3) | 34.8 (±10.5) | 34.8 (±10.3) | | | | |

SD standard deviation
| Authors and year | Domain 1—objective(s) and subject characteristics | Domain 2—study design | Domain 3—methodology characterisation | Domain 4—descriptive anatomy |
|------------------|-----------------------------------------------|-------------------------|---------------------------------------|-----------------------------|
| Amabile et al. (2016) [56] | Y Y Y | Y Y Y Y | Y | Y N Y Y Y |
| Bakouny et al. (2017) [37] | Y Y Y | Y Y Y Y | Y N Y N Y |
| Bassani et al. (2019) [20] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Endo et al. (2014) [46] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Endo et al. (2016) [40] | Y Y Y | Y Y Y Y | Y N Y Y Y |
| Gongnet et al. (2006) [54] | Y Y Y | Y Y Y Y | Y N Y Y Y |
| Gelb et al. (1995) [7] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Granito et al. (2014) [48] | Y Y Y | Y Y Y Y | Y Y Y N Y |
| Hammerberg et al. (2003) [61] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Hasegawa et al. (2016) [51] | Y Y Y | Y Y Y Y | Y N Y N Y |
| Hasegawa et al. (2017) [53] | Y Y Y | Y Y Y Y | Y N Y Y Y |
| Himmen (2004) [47] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Hu et al. (2016) [42] | Y Y Y | Y Y Y Y | Y N Y N Y |
| Hu et al. (2020) [32] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Iyer et al. (2016) [6] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Janssen et al. (2009) [36] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Kim et al. (2014) [31] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Korovessis et al. (1998) [44] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Ladjé et al. (2019) [53] | Y Y Y | Y Y Y Y | Y N Y N Y |
| Lee et al. (2015) [49] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Le Huec et al. (2016) [38] | Y Y Y | Y Y Y Y | Y N Y N Y |
| Lee et al. (2015) [59] | Y Y Y | Y Y Y Y | Y Y Y N Y |
| Park et al. (2013) [58] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Pavlovic et al. (2013) [52] | Y Y Y | Y Y Y Y | Y N Y N Y |
| Schwab et al. (2006) [50] | Y Y Y | Y Y Y Y | Y N Y Y Y |
| Sudhir et al. (2016) [39] | Y Y Y | Y Y Y Y | Y Y Y N Y |
| Uehara et al. (2019) [41] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Urrutia et al. (2014) [60] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Valle et al. (2005) [57] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Yang et al. (2017) [37] | Y Y Y | Y Y Y Y | Y Y Y N Y |
| Yeh et al. (2018) [35] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Yokoyama et al. (2017) [33] | Y Y Y | Y Y Y Y | Y Y Y N Y |
| Yukawa et al. (2018) [43] | Y Y Y | Y Y Y Y | Y N Y N Y |
| Zhu et al. (2014) [45] | Y Y Y | Y Y Y Y | Y Y Y Y Y |
| Authors and year | Q1 | Q2 | Q3 | Q4 | Domain 5—reporting of results | Q1 | Q2 | Q3 | Q4 | Domain results | 1 | 2 | 3 | 4 | 5 | Overall quality |
|------------------|----|----|----|----|-------------------------------|----|----|----|----|----------------|----|----|----|----|----|----------------|
| Amabile et al. (2016) [56] | Y | Y | Y | Y | Domain 5—reporting of results | Y | Y | N | Y | Domain results | L | L | H | L | H | Overall quality | M |
| Balkony et al. (2017) [37] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Bassani et al. (2019) [20] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | H | M |
| Endo et al. (2014) [46] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Endo et al. (2016) [40] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Gongnet et al. (2006) [54] | Y | Y | Y | Y | Y | Y | Y | N | Y | L | L | H | L | L | M |
| Gelp et al. (1995) [7] | N | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | H | L | M |
| Granito et al. (2014) [48] | Y | Y | Y | Y | Y | Y | N | Y | Y | L | L | H | L | H | M |
| Hammerberg et al. (2003) [61] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Hasegawa et al. (2016) [51] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Hasegawa et al. (2017) [53] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | M |
| Hinman (2004) [47] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Hu et al. (2016) [42] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | M |
| Hu et al. (2020) [32] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Iyer et al. (2016) [6] | Y | Y | N | N | Y | Y | Y | N | Y | L | L | L | H | H | M |
| Janssen et al. (2009) [36] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Kim et al. (2014) [31] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Koevessis et al. (1998) [44] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Lafoge et al. (2019) [53] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Lee et al. (2015) [49] | Y | Y | Y | Y | Y | Y | Y | N | Y | L | L | L | L | H | M |
| Le Huec et al. (2016) [38] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Oe et al. (2015) [59] | N | Y | N | N | Y | Y | Y | N | Y | L | L | H | L | L | M |
| Park et al. (2013) [58] | Y | Y | Y | Y | Y | Y | Y | N | Y | L | L | L | L | L | H |
| Pavlovic et al. (2013) [52] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Schwab et al. (2006) [50] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Sudhir et al. (2016) [39] | N | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | H | L | M |
| Uehara et al. (2019) [41] | N | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | H | L | M |
| Urrutia et al. (2014) [60] | N | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | M |
| Vallee et al. (2005) [57] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Yang et al. (2017) [37] | Y | Y | N | Y | Y | Y | Y | Y | Y | L | L | H | H | L | M |
| Yeh et al. (2018) [35] | Y | Y | N | Y | Y | Y | Y | Y | Y | L | L | L | H | L | M |
| Yokoya et al. (2017) [33] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | H | L | L | M |
| Yukawa et al. (2018) [43] | Y | Y | Y | Y | Y | Y | N | Y | Y | L | L | H | L | H | M |
| Zhu et al. (2014) [45] | Y | Y | Y | Y | Y | Y | Y | Y | Y | L | L | L | L | L | H |
| Body reference | Correlation (p-value) | Sample size | Age group | Mean (SD) (degrees) | Range (degrees) | Group comparisons (p-value) |
|----------------|-----------------------|-------------|-----------|---------------------|----------------|---------------------------|
| All ethnicity  | 0.23 (0.21)           | 2154        | x < 40    | 39.8 (±5.87)        | 33.93–45.67   | x < 40 – x > 60 (0.015)   |
|                |                       |             | 40 < x < 60| 38.54 (±4.68)      | 33.86–43.22   | 40 < x < 60 – x > 60 (0.32) |
|                |                       |             | x > 60    | 43.56 (±10.5)      | 33.08–54.04   | x < 50 – x > 50 (0.17)    |
|                |                       |             | x < 50    | 34.94 (±4.58)      | 33.76–44.12   | x < 40 – x > 60 (0.049)   |
|                |                       |             | x > 50    | 43.11 (±9.56)      | 33.55–52.67   | 40 < x < 60 – x > 60 (0.49) |
| All genders    | 0.64 (0.12)           | 212         | x < 40    | 44.53 (±4.16)      | 40.37–48.17   | x < 40 – x > 60 (0.49)    |
|                |                       |             | 40 < x < 60| 47.7 (±4.53)       | 43.17–52.23   | 40 < x < 60 – x > 60 (0.49) |
|                |                       |             | x > 60    | 48.75 (±1.48)      | 47.27–50.23   | x < 50 – x > 50 (0.07)     |
|                |                       |             | x < 50    | 44.53 (±3.4)       | 41.13–47.92   | x < 40 – x > 60 (0.015)   |
|                |                       |             | x > 50    | 49.47 (±1.63)      | 47.84–51.09   | x < 40 – x > 60 (0.015)   |
| T2–T12         | 0.45* (0.048)         | 1617        | x < 40    | 33.88 (±4.88)      | 29–38.77      | x < 40 – x > 60 (0.00052) |
|                |                       |             | 40 < x < 60| 40.1 (±2.69)       | 37.41–42.79   | x < 40 – x < 60 (0.055)   |
|                |                       |             | x > 60    | 47.63 (±7.33)      | 40.29–54.96   | 40 < x < 60 – x > 60 (0.16) |
|                |                       |             | x < 50    | 34.11 (±4.77)      | 29.34–38.87   | x < 50 – x > 50 (0.0003)  |
|                |                       |             | x > 50    | 45.63 (±6.47)      | 39.17–52.1    | x < 50 – x > 50 (0.0003)  |
| T4–T12         | 0.87 a (0.02)         | 120         | x < 40    | 42.2 (±1.41)       | 40.79–43.61   | x < 40 – x > 60 (0.00011) |
|                |                       |             | 40 < x < 60| 47.7 (±4.52)       | 43.17–52.23   | x < 40 – x < 60 (0.036)   |
|                |                       |             | x > 60    | 48.75 (±1.48)      | 47.27–50.23   | x < 40 – x > 60 (0.036)   |
|                |                       |             | x < 50    | 42.97 (±1.66)      | 41.3–44.63    | x < 40 – x < 60 (0.036)   |
|                |                       |             | x > 50    | 49.47 (±1.63)      | 47.84–51.09   | x < 40 – x < 60 (0.036)   |
| T5–T12         | 0.82 (0.18)           | 190         | x < 50    | 37.5 (±0.7)        | 36.79–38.21   | x < 50* (0.043)           |
|                |                       |             | x > 50    | 42.75 (±1.77)      | 40.98–44.52   | x < 50* (0.043)           |
| T5–T12         | 0.65* (0.03)          | 339         | x < 40    | 30.05 (±1.63)      | 28.42–31.68   | x < 50* (0.043)           |
|                |                       |             | 40 < x < 60| 35.02 (±2.62)      | 32.4–37.64    | x < 50* (0.043)           |
|                |                       |             | x > 60    | 35.93 (±2.19)      | 33.73–38.12   | x < 50* (0.043)           |
|                |                       |             | x < 50    | 32.2 (±2.98)       | 29.22–35.18   | x < 50* (0.043)           |
|                |                       |             | x > 50    | 35.73 (±2.42)      | 33.31–38.15   | x < 50* (0.043)           |
| T5–T12         | 0.72* (0.0000006)     | 2919        | x < 40    | 25.39 (±4.33)      | 21.06–29.71   | x < 40 (0.19)             |
|                |                       |             | 40 < x < 60| 27.69 (±2.26)      | 25.43–29.95   | x < 40 (0.19)             |
|                |                       |             | x > 60    | 31.76 (±3.49)      | 28.27–35.24   | x > 60* (0.033)           |
Table 4 Correlation analysis, normative values and between-group difference (Continued)

| Body reference | Correlation (p-value) | Sample size | Age group | Mean (SD) (degrees) | Range (degrees) | Group comparisons (p-value) |
|----------------|-----------------------|-------------|-----------|--------------------|----------------|-----------------------------|
|                |                       |             |           |                    |                |                             |
|                | 1079 x < 50 25.55 (±3.86) 21.69–29.41 | x < 50* (0.0097) |
|                | 1486 x > 50 30.81 (±3.58) 27.23–34.39 | x > 50* (0.0017) |
| Europe T1–T12 | 0.8 (0.57) 404 250 x < 50 44.87 (±3.96) 40.9–48.83 | 154 x > 50 55.7 (±7.04) 48.66–62.74 |
|                | 0.88 (0.00076) 860 610 x < 50 36.18 (±2.41) 33.77–38.6 | 253 x > 50 47.08 (±7.77) 39.31–54.84 |
| Europe T4–T12 | 0.8 (0.57) 301 x < 40 36.16 (±0.39) 35.77–36.55 | 145 x > 60 33.57 (±4.92) 28.65–38.49 |
|                | – 0.17 (0.55) 724 278 x < 40 44.87 (±3.63) 31.06–43.83 | 301 40 < x < 60 36.16 (±0.39) 35.77–36.55 |
|                | 0.65 (0.0024) 1254 383 x < 40 45.9 (±7.78) 38.12–53.68 x > 60 (0.19) |
|                | 0.6 (0.0015) 126 73 x < 40 44.33 (±4.16) 40.17–48.49 40 < x < 60 (0.27) |
|                | 28 40 < x < 60 45.9 (±7.78) 38.12–53.68 |
|                | 25 x > 60 49.45 (±6.76) 42.73–56.17 x < 50 (0.41) |
|                | 88 x < 50 43.35 (±3.92) 39.43–47.27 x > 50 (0.52) |
|                | 38 x > 50 50.1 (±4.88) 45.22–54.98 |
| Gender Female  | T1–T12 | 1254 383 x < 40 27.27 (±5.27) 21.99–32.54 x > 60 (0.19) |
|                | 0.65 (0.00024) 687 x < 60 33.45 (±5.18) 28.27–38.63 x > 60 (0.19) |
|                | 440 x < 50 27.49 (±4.5) 22.98–31.99 x < 40 (0.92) |
|                | 814 x > 60 32.71 (±5.12) 27.58–37.83 40 < x < 60 (0.41) |
|                | 184 40 < x < 60 29.82 (±4.16) 25.65–33.98 T2–T12 |
| Male           | T1–T12 | 0.03 (0.94) 571 172 x < 40 40.38 (±7.22) 33.15–47.6 x > 60 (0.97) |
|                | 0.07 (0.88) 86 59 x < 40 44.8 (±6.87) 37.93–51.67 TS–T12 |
|                | 274 40 < x < 60 38.7 (±3.74) 34.96–42.43 x < 50 (0.35) |
|                | 125 x > 60 38.57 (±3.16) 35.41–41.72 x > 50 (0.75) |
|                | 390 x < 50 39.56 (±5.83) 33.72–45.39 |
|                | 181 x > 50 38.78 (±2.61) 37.47–41.39 |
|                | T2–T12 | 0.07 (0.88) 86 59 x < 40 44.8 (±6.87) 37.93–51.67 TS–T12 |
|                | 20 x > 60 49.2 (±4.53) 44.67–53.73 x < 40 (0.69) |
|                | 63 x < 50 48.81 (±9.29) 39.21–57.79 40 < x < 60 (0.62) |
|                | 23 x > 50 48.97 (±3.23) 45.74–52.19 x > 60 (0.4) |
|                | T5–T12 | 0.51* (0.0085) 1393 540 x < 40 26.17 (±4.9) 21.28–31.07 x < 50 (0.81) |
|                | 152 40 < x < 60 31.87 (±8.88) 22.99–40.75 x > 50 (0.57) |
|                | 701 x > 60 32.28 (±3.66) 28.62–35.99 |
|                | 584 x < 50 28.27 (±7.7) 20.55–35.99 |
|                | 809 x > 50 31.71 (±4.46) 27.24–36.17 |

x < 40 people younger than 40 years old, 40 < x < 60 people between 40 and 60 years old, x > 60 people older than 60 years old, x < 50 people younger than 50 years old, x > 50 people older than 50 years old, T thoracic vertebra

*Statistical significance for p < 0.05 (t-test)

Statistical significance for p < 0.0167 (ANOVA with Bonferroni post hoc correction)
technical difficulties with visualising the vertebrae above T4 from lateral radiographs [2], measuring kyphosis from T5 may provide more accurate measurements.

**Normative values**

The normative values surpassed 40° in 65% of the analysis. This finding challenges the accuracy of the current threshold used for defining normality (i.e. 40°). This cut-off was first introduced by Roaf in 1960 [1], but without supporting evidence for it. Despite subsequent studies showing that healthy children, adolescents and adults could have thoracic curvatures exceeding 40° [6, 65], this value is still used in practice [3, 4]. Some authors suggested moving this cut-off to 50° [2]. However, even this suggestion may not decrease the chances of misclassifying patients, since 35% of the ranges presented in this review surpassed 50°. Using a range of 20–60° [9] may seem more appropriate, since the ranges provided never exceeded 60°. Nonetheless, people x < 40 appeared to have a significantly smaller kyphosis than those x > 60. Consequently, using the same reference values for both groups may lead to misclassification anyway. When kyphosis was measured between T4/5 and T12, its value significantly differed also between people x < 50 and x > 50. This may indicate a higher measurement precision when those body references were used. Thoracic kyphosis varied depending on the body references selected to calculate it, with a trend showing that including higher vertebrae leads to greater values. Therefore, using specific reference values, like those presented in this review, which account for age and body references, could be the most accurate alternative for clinicians.

**Gender and ethnic group differences**

Thoracic kyphosis does not seem to be influenced by gender, since the between-group mean difference never reached statistical significance. Although the precision of the results could have been affected by the small number of studies subdividing their sample by age groups and gender, these findings align with previous evidence [7, 57].

Significant differences in kyphosis between the ethnic groups were seen, with Europeans and North Americans showing a greater kyphosis than Asians. Genetic differences may explain this result. A twins study found that thoracic kyphosis is influenced by genetics and that it also negatively correlates with bone mineral density [66], also related to genetics [67]. However, other lifestyle factors, such as sports, could also influence thoracic curvature [68], but no data were available to investigate those relationships. Since only 14 studies specified the sample ethnicity [20, 32, 34–38, 41–46, 59], people were grouped according to geography. This can represent a limitation since some areas have inhabitants from different socio-cultural backgrounds. Most of the studies that specified sample ethnicity included people from Asia [32, 34, 35, 37, 38, 41–43, 45, 59] or Europe [20, 36, 38, 44], which further affects the reliability of the results for North America.

**Strengths and limitations**

This review employed rigorous methods, with transparent reporting (PRISMA and SWiM guidelines), and a completed PRISMA checklist relative to this article can be found in Additional file 3. The main strength of this review lies in the high quality of studies included and the large sample size utilised for computing the values presented. These factors strengthen the confidence in study findings. No information about kyphosis measured with a kyphometer or flexicurve was provided because of poor information retrieval, perhaps due to the limited sensitivity of the search tool [15]. The AQUA tool was utilised to assess study quality, but data regarding its validity and reliability is lacking [17]. Since clinical and methodological heterogeneity can preclude a meta-analysis [25], and concerns regarding the reliability of
Table 5 Synthesis of results

| Sample size | Number of studies | Results bias | Inconsistency (-) Consistency (+) | Imprecision (-) Precision (+) | Indirectness | Study limitations/quality | Effect size/dose response | Possible confounding/bias | Overall quality |
|-------------|------------------|--------------|-----------------------------------|-------------------------------|--------------|--------------------------|--------------------------|---------------------------|------------------|
| Correlation Kyphosis and age | 6793 | 30 | Table 4 ND | (+) All correlations had positive sign | T1–T12: p > 0.05 (-) T2–T12: p > 0.05 (-) T4–T12: p < 0.05 (+) T5–T12: p < 0.05 (+) | None | (+) 9 HQ, 21 MQ | N/A | (+)(+)+(+)| Moderate |
| Reference values | 6793 | 30 | Table 4 ND | (−) Statistical significance was achieved in 6 out of 16 cases | (−) Ranges did not overlap in 2 out of 6 cases. The between-range difference was always smaller than the SEM. | None | (+) 9 HQ, 21 MQ | N/A | (+) 12 articles divided sample by age groups | (+) (+) Low |
| Ethnic group differences | 5788 | 25 | Table 4 ND | (−) Statistical significance was achieved in 6 out of 9 cases. | (−) Ranges did not overlap in 2 out of 6 cases. In 1 of the 2 cases, the between-range difference was greater than the SEM. | None | (+) 8 HQ, 17 MQ | N/A | (+) 14 articles specified the ethnicity of their sample | (+)(+)+ Low |
| Gender differences | 3937 | 13 | Table 4 ND | (+) No between-group comparisons achieved statistical significance | (−) | None | (+) 3 HQ, 10 MQ | N/A | (+)(+)+(+)| Moderate |

ND not detected, HQ high-quality, MQ moderate quality, N/A not applicable, T thoracic vertebra, SEM standard error of measurement
the results of the meta-analysis carried out on observational studies exist [69], the authors considered a narrative synthesis most appropriate. Finally, the sample utilised to create the normative values presented was not randomly selected from the general population, but it was created by combining the samples of the individual studies included in the review, and this could represent a form of selection bias. However, the rigorous methodology employed, the size and the heterogeneity of the sample may partially mitigate this limitation.

Clinical implications
Surgical interventions aiming to correct adult spinal deformities are recommended in those cases with progressive deformities, significant neural compromising, pain or functional limitations, and that did not respond to conservative management [9]. To help these patients, different surgical approaches are available, from minimally invasive operations, such as laminectomies, to deformity correction and vertebral fusion surgeries. These more invasive interventions may target only a limited and specific number of vertebrae in mild and moderate cases or extensive portions of the thoracic and lumbar spine in more severe cases [70], reaching as high as T3–T4 in some instances [71]. These more invasive interventions are associated with high risk of complications and worse functional outcomes if the surgical correction is suboptimal; thus, careful surgical planning is paramount [70]. Among the individual patient’s characteristics to be considered when planning for surgery, there are patient’s age [72] and ethnicity [71]; consequently, we believe that the normative values provided in this review, which account specifically for these characteristics, despite being supported by low-quality evidence, may prove beneficial in a clinical context. This information may help clinicians deciding and planning their interventions.

Conclusion
This review provides evidence that a positive correlation between kyphosis and age exists. It also shows that thoracic kyphosis seems to not be influenced by gender, but to vary depending on ethnicity, age, and the body references used to measure it. The normative values of kyphosis currently used in clinical practice may not reduce the chances of misclassifying patients, since they do not account for those characteristics, and they may not be precise enough to correctly inform clinicians when planning and performing corrective spinal surgeries. Therefore, using specific reference values, such as those presented in this study, which account for body reference, age, and ethnicity, when assessing and treating patients may represent the most accurate solution for clinicians.

Abbreviations
ANOVA: Analysis of variance; AQUA: Anatomical Quality Assessment; C: Cervical vertebra; GRADE: Grading of Recommendations, Assessment, Development and Evaluation; H: High; HQ: High-quality; L: Low; M: Moderate; MQ: Moderate quality; N: No; N/A: Not applicable; ND: Not detected; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for protocols; Q: Question; SD: Standard deviation; SEM: Standard error of measurement; SPIDER: Sample, Phenomenon of Interest, Design, Evaluation, Research type; SWIM: Synthesis Without Meta-analysis; T: Thoracic vertebra; x < 40: People younger than 40 years old; x > 60: People older than 60 years old; x < 50: People younger than 50 years old; x > 50: People older than 50 years old; Y: Yes; 40 < x < 60: People between 40 and 60 years old

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s13018-021-02592-2.

Additional file 1: Table S1. Examples of search strategy.
Additional file 2: Table S2. Supplementary table for the AQUA tool.
Additional file 3: PRISMA 2020 Checklist.

Acknowledgements
N/A.

Authors’ contributions
All authors conceptualised and designed the manuscript. MZ led the protocol development supported by all authors. SL helped MZ in the screening, study selection, risk of bias assessment and data extraction. MZ drafted the initial manuscript with NRH. SL provided guidance on the design, topic, methodology and analyses. All authors reviewed and commented on each draft of the manuscript. All authors have approved and contributed to the final manuscript.

Funding
No funds received in support of this work.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations
Ethics approval and consent to participate
N/A.

Consent for publication
N/A.

Competing interests
The authors declare that they have no competing interests.

Author details
1 Physiotherapy Department, St John & St Elizabeth Hospital, 60 Grove End Rd., St John’s Wood, London, UK. 2 Bermuda Hospitals Board, King Edward Memorial Hospital, 7 Point Finger Road, Paget, DV 04, Bermuda. 3 Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, College of Life and Environmental Sciences, University of Birmingham, Birmingham, UK.

Received: 18 February 2021 Accepted: 28 June 2021
Published online: 09 July 2021

References
1. Roaf R. Vertebral growth and its mechanical control. J Bone Joint Surg Br. 1960;42-B(1):40–59. https://doi.org/10.1302/0301-620X.42B1.40.
healthy adults. J Orthop Surg. 2016;24(1):92–6. https://doi.org/10.11177/23094990160200121.

41. Uehara M, Takahashi J, Ikegami S, Tokuoka R, Nishimura H, Sakai N, et al. Sagittal spinal alignment deviation in the general elderly population: a Japanese cohort survey randomly sampled from a basic resident registry. Spine J. 2019;19(2):349–56. https://doi.org/10.1016/j.spinee.2018.06.346.

42. Hu P, Yu M, Sun Z, Li W, Jiang L, Wei F, et al. Analysis of global sagittal postural patterns in asymptomatic Chinese adults. Asian Spine J. 2016;10(2):282–8. https://doi.org/10.1016/j.asj.2016.10.2.282.

43. Yukawa Y, Kato F, Suda K, Yamagata M, Ueta T, Yoshida M and Yoshida M. Normative data for parameters of sagittal spinal alignment in healthy subjects: an analysis of gender specific differences and changes with aging in 625 asymptomatic individuals. Eur Spine J. 2018;27(2):426–32. https://dx.doi.org/10.1007/s00586-016-4807-7.

44. Korovessis PG, Stamatakis MV, Baikousis AG. Reciprocal angulation of thoracic kyphosis and postural stiffness in healthy adults. Eur Spine J. 2004;13(4):411–7. https://doi.org/10.1007/s00586-004-0022.

45. Granito RN, Aveiro MC, Rennó ACM, et al. Degree of thoracic kyphosis and peak torque of trunk flexors and extensors among healthy women. Rev Bras Ortop. 2011;46(3):286–91. https://doi.org/10.1590/S1807-12682011000300002.

46. Endo K, Suzuki H, Nishimura H, Tanaka H, Shishido T, Yamamoto K. Characteristics of sagittal spine-pelvic alignment in Japanese young adults. Asian Spine J. 2018;12(5):599–604. https://doi.org/10.1016/j.asj.2018.04.099.

47. Hinman MR. Comparison of thoracic kyphosis and postural stiffness in younger and older women. Spine J. 2004;4(4):413–7. https://doi.org/10.1016/j.sj.2004.01.002.

48. Garito RN, Aveiro MC, Rennó ACM, et al. Degree of thoracic kyphosis and peak torque of trunk flexors and extensors among healthy women. Rev Bras Ortop. 2011;46(3):286–91. https://doi.org/10.1590/S1807-12682011000300002.

49. Lee S, Son E, Seo E, et al. Factors determining cervical spine sagittal balance: normative data on 100 asymptomatic subjects. Spine. 2013;38(8):967–72. https://doi.org/10.1097/BRS.0b013e31828802c2.

50. Schwab F, Lafage V, Boyce R, Skalli W, Farcy JP. Gravity line analysis in adult spine. Spine. 2004;29(14):1711–7. https://doi.org/10.1097/01.brs.0000012274.38382.06.

51. Boeker EH, Moe JH, Winter RB, Koop SE. Determination of “normal” thoracic kyphosis: a roentgenographic study of 121 “normal” children. J Pediatr Orthop. 2000;20(6):796–8. https://doi.org/10.1097/00004694-200011000-00001.

52. Stone MA, Osei-Bordom D, Inman RD, et al. Heritability of spinal curvature and its relationship to disc degeneration and bone mineral density in female adult twins. Eur Spine J. 2014;23(11):2387–94. https://doi.org/10.1007/s00586-014-3477-6.

53. Hasegawa K, Okamoto M, Hatsukiko S, Shimoda H, Ono M, Watanabe K. Surgical treatment for adult spinal deformity: conceptual alignment targets for adults spinal deformity correction depend on pelvic alignment. Asian Spine J. 2020;14(6):926–32. https://doi.org/10.1007/s12910-020-0251-2.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.