Heritability of sleep duration and quality: A systematic review and meta-analysis

Desana Kocevska a, *, Nicola L. Barclay b, Wichor M. Bramer c, Philip R. Gehrman d, Eus J.W. Van Someren a, e, f

a Department of Sleep and Cognition, Netherlands Institute for Neuroscience, An Institute of the Royal Netherlands Society for Arts and Sciences, Amsterdam, The Netherlands
b Sleep and Circadian Neuroscience Institute (SCNI), Department of Clinical Neurosciences, University of Oxford, UK
c Medical Library, Erasmus MC University Medical Center, Rotterdam, The Netherlands
d Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA
e Department of Integrative Neurophysiology and Psychiatry, VU University, Amsterdam UMC, Amsterdam, The Netherlands
f Center for Neurogenomics and Cognitive Research, VU University, Amsterdam UMC, Amsterdam, The Netherlands

SUMMARY

Epidemiological and interventional research has highlighted sleep as a potentially modifiable risk factor associated with poor physical and mental health. Emerging evidence from (behavioral) genetic research also shows that sleep characteristics are under strong genetic control. With this study we aimed to meta-analyze the literature in this area to quantify the heritability of sleep duration and sleep quality in the general population. We conducted a systematic literature search in five online databases on January 24th 2020. Two authors independently screened 5644 abstracts, and 160 complete articles for the inclusion criteria of twin studies from the general population reporting heritability statistics on sleep duration and/or quality, and written in English. We ultimately included 23 papers (19 independent samples: 45,328 twins between 6 mo and 88 y) for sleep duration, and 13 papers (10 independent samples: 39,020 twins between 16 and 95 y) for sleep quality. Collectively, we showed that 46% of the variability in sleep duration and 44% of the variability in sleep quality is genetically determined. The remaining variation in the sleep characteristics can mostly be attributed to the unique environment the twins experience, although the shared environment seemed to play a role for the variability of childhood sleep duration. Meta-analyzed heritability estimates for sleep duration, however, varied substantially with age (17% infancy, 20–52% childhood, 69% adolescence and 42–45% adulthood) and reporter (8% parent-report, 38–52% self-report). Heritability estimates for actigraphic and Polysomnography (PSG)-estimated sleep were based on few small samples, warranting more research. Our findings highlight the importance of considering genetic influences when aiming to understand the underlying mechanisms contributing to the trajectories of sleep patterns across the lifespan.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

Ample epidemiological evidence points towards an association between poor sleep (insufficient sleep duration or quality) and worse health outcomes [1]. Intervention research has also emphasized that sleep characteristics should be recognized as potentially modifiable determinants of health and well-being [2]. Despite sleep still being considered a lifestyle characteristic [3], emerging evidence shows that sleep duration and quality are, to some extent, genetically inherited traits.

Twin studies have shown that a substantial amount of variability in sleep characteristics are genetically determined [4]. This has been supported by a number of candidate gene studies as well as genome-wide association studies (GWAS) [5–7]. However, whilst the molecular genetic studies provide important insights into the role common variants play in numerous traits [5,6,8], such studies have only identified an insubstantial number of genes that together account for a small proportion of variability in sleep duration, and thus missing heritability remains an issue [9].

* Corresponding author. Meibergdreef 47, 1105 BA Amsterdam, The Netherlands.
E-mail address: d.kocevska@nin.knaw.nl (D. Kocevska).
Numerous twin studies (which despite concerns [10,11] remain the ‘gold standard’ for estimating heritability [12–14]), have provided heritability estimates for both sleep duration and sleep quality. However, considerable variation in the estimates can also be noted. Specifically, multiple twin studies from 1983 onwards have estimated heritability of sleep duration in various twin-samples from the general population, with estimates ranging between 0% [15] to 71% [16]. Not only the duration of sleep, but also self-reported sleep quality has a substantial genetic component, with twin-study heritability estimates ranging from 24% [17] to 53% [18]. Given quality has a substantial genetic component, with twin-study heritability estimates ranging from 71% [16]. Not only the duration of sleep, but also self-reported sleep quality (31%) by meta-analyzing studies that partially overlap with those we include [19]. However, this study based their meta-estimates on: 1) estimates from structural equation twin model-fitting approaches, which highly depends on model choice and sample size; 2) studies investigating populations older than 6 y, thus excluding young children.

The aim of the present study is thus to systematically meta-analyze studies investigating the heritability of sleep duration and sleep quality across the lifespan and to perform a meta-analysis of these studies to generate a robust estimate of the genetic contribution to individual differences in sleep characteristics. Additionally, this study aims to determine whether the heritability of sleep duration and quality is moderated by sex, age of sample, or assessment method.

Methods

Literature search

We used a standardized systematic literature searching [20] and reviewing [21] protocol. The following databases were searched to identify relevant articles: Embase.com, Medline via Ovid, Web of Science core Collection, Cochrane CENTRAL Register of Trials via Wiley and Google Scholar [22]. The searches were designed by an experienced information specialist (WMB). The initial search was conducted on February 20th 2019, which was later updated to search for additional items on 24th January 2020, with a search strategy outlined in the Supplementary Text.

Study selection procedure

The following criteria were used to select studies:

Inclusion criteria

A primary research study that:

1) investigated subjective sleep quality or sleep duration assessed with questionnaires (parent- or self-reported), actigraphy, sleep diaries or polysomnography
2) is a general population sample
3) is a behavioral genetic study utilizing one of the following designs: classical twin study, twin/sibling study
4) reports descriptive information about sample (n, male/female ratio)
5) reports statistics necessary for effect size calculations (intraclass correlation coefficients; variance components)
6) reported on independent samples, or different data from overlapping samples
7) is published in English

Exclusion criteria

1) population with psychiatric or medical disorder
2) studies with only monozygotic twins available
3) studies with no heritability estimates or twin correlations
4) reviews
5) meta-analyses

Our initial search yielded 7541 hits, after which 5040 remained once duplicates were removed. Our updated search identified yielded a further 727 hits after duplicates were removed. The first two authors independently screened title and abstracts, and in a next step the full-texts were independently screened for eligibility by the two authors. Differences in decision between the two authors who read the papers were resolved by further discussion. Reference sections of those included were assessed, and no additional papers were identified.

Data extraction

The following data were extracted from each study using standardized coding sheets:

- Publication Date, Authors, Title, Country, Registry name, Total sample size, Subgroup sample sizes (males, females, monozygotic (MZ), dizygotic (DZ)), Study type (classical twin study, twin/sibling study), Measurement of Sleep Duration and Sleep Quality (self-reported, parent reported, actigraphy, Electroencephalogram (EEG)), twin intraclass correlations (rMZ and rDZ), sampling variances for MZ and DZ, variance components (A, C and E) and Sample sex and age (coded as 1 = infancy < 1 y, 2 = early childhood 1–3 y, 3 = middle childhood 4–8 y, 4 = older childhood 9–12 y, 5 = adolescence, 6 = adulthood, 7 = older adulthood). Several studies provided more than one effect size (e.g., separately for sex, age group, reporter or assessment method). Tables 1 and 2 provide an overview of the included studies for sleep duration and quality, respectively.

The twin design

In quantitative behavioral genetics, the classical twin design relies on knowledge of the relative differences in genetic and environmental correlations between MZ and DZ twins [13]. This enables parsing out variance in a phenotype of interest into relative proportions of additive genetic (where genes “add up” to influence behavior; A), shared environmental (C) and non-shared environmental (E) influences. MZ twins share on average 100% of their segregating genes, whilst DZ twins share around 50% of their segregating genes. Thus, additive genetic influences can be assumed to play a role in a phenotype if the MZ twin correlation (rMZ) is greater than the DZ twin correlation (rDZ). The shared environment is estimated to be equal between both MZ and DZ twins and is typically attributed to family or social environments that are shared between twins within a family that account for their similarity, and thus are equated at one for both MZ and DZ twins. Non-shared environmental influences on the other hand account for unique experiences of twins which contribute to their differences. Thus, non-shared environmental influences are estimated to be equal between pairs of both MZ and DZ twins and is typically attributed to family or social environments that are shared between twins within a family that account for their similarity, and thus are equated at one for both MZ and DZ twins. Non-shared environmental influences on the other hand account for unique experiences of twins which contribute to their differences. Thus, non-shared environmental influences are estimated to be equal between pairs of both MZ and DZ twins. Using these correlations we can calculate a measure of heritability, denoted h2, using Falconer’s formula as follows: A = 2 (rMZ−rDZ). The proportion of shared environmental influences contributing to a phenotype is calculated as: C = rMZ−A. Finally, non-shared environmental influences on a phenotype are the only factors that account for differences between identical twins, hence: E = 1−rMZ.
Table 1
Characteristics of twin studies on sleep duration.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>nMZ</th>
<th>nDZ</th>
<th>Age, years (range)</th>
<th>Method</th>
<th>rMZ</th>
<th>rDZ</th>
<th>h²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barclay [15]</td>
<td>2010</td>
<td>UK</td>
<td>380</td>
<td>654</td>
<td>20 (18–27)</td>
<td>Self-report</td>
<td>0.23</td>
<td>0.28</td>
<td>0.71</td>
</tr>
<tr>
<td>Breitenstein [40]</td>
<td>2018</td>
<td>USA</td>
<td>178</td>
<td>234</td>
<td>8.5</td>
<td>Parent-report</td>
<td>0.87</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Actigraphy</td>
<td>0.84</td>
<td>0.47</td>
<td>0.23</td>
</tr>
<tr>
<td>Brettinestone [34]</td>
<td>2018</td>
<td>USA</td>
<td>151</td>
<td>210</td>
<td>1–5</td>
<td>Parent-report</td>
<td>0.86–0.96</td>
<td>0.79–0.92</td>
<td>0.15–0.36</td>
</tr>
<tr>
<td>Brescianini [52]</td>
<td>2011</td>
<td>Italy</td>
<td>254</td>
<td>374</td>
<td>1.3</td>
<td>Parent-report</td>
<td>0.95</td>
<td>0.81</td>
<td>0.31</td>
</tr>
<tr>
<td>Butovik [45]</td>
<td>2014</td>
<td>Croatia</td>
<td>210</td>
<td>468</td>
<td>15–22</td>
<td>Self-report</td>
<td>0.62</td>
<td>0.20</td>
<td>0.16</td>
</tr>
<tr>
<td>De Castro [53]</td>
<td>2002</td>
<td>US</td>
<td>172</td>
<td>258</td>
<td>41.9 ± 10.2</td>
<td>Diary</td>
<td>0.30</td>
<td>0.04</td>
<td>0.30</td>
</tr>
<tr>
<td>Fisher [37]</td>
<td>2012</td>
<td>UK</td>
<td>1190</td>
<td>2540</td>
<td>1.3 (1.2–2.3)</td>
<td>Parent-report</td>
<td>0.92</td>
<td>0.80</td>
<td>0.26</td>
</tr>
<tr>
<td>Gehrmann [28]</td>
<td>2019</td>
<td>USA</td>
<td>90</td>
<td>100</td>
<td>16–40</td>
<td>Actigraphy</td>
<td>0.59</td>
<td>0.15</td>
<td>0.49</td>
</tr>
<tr>
<td>Giedda &amp; Breen [27]</td>
<td>1979</td>
<td>Italy</td>
<td>154</td>
<td>152</td>
<td>6–18</td>
<td>Self-report</td>
<td>0.71–0.97</td>
<td>0.71–0.85</td>
<td></td>
</tr>
<tr>
<td>Gregory [16]</td>
<td>2006</td>
<td>UK</td>
<td>200</td>
<td>398</td>
<td>8 (8.2–8.9)</td>
<td>Parent-report</td>
<td>0.76</td>
<td>0.36</td>
<td>0.71</td>
</tr>
<tr>
<td>Heath [41]</td>
<td>1990</td>
<td>Australia</td>
<td>3584</td>
<td>2200</td>
<td>17–88</td>
<td>Self-report</td>
<td>0.39–0.41</td>
<td>0.09–0.24</td>
<td>0.09</td>
</tr>
<tr>
<td>Hublin [42]</td>
<td>2013</td>
<td>Finland</td>
<td>3938</td>
<td>8390</td>
<td>&gt;18</td>
<td>Self-report</td>
<td>0.21–0.44</td>
<td>0.01–0.26</td>
<td>0.30–0.32</td>
</tr>
<tr>
<td>Inkerkum [29]</td>
<td>2018</td>
<td>Switzerland</td>
<td>32</td>
<td>20</td>
<td>12.7 ± 1</td>
<td>Actigraphy</td>
<td>0.55–0.69</td>
<td>–0.17–0.57</td>
<td>0.15–0.68</td>
</tr>
<tr>
<td>Liu [47]</td>
<td>2012</td>
<td>China</td>
<td>664</td>
<td>284</td>
<td>21–72</td>
<td>Self-report</td>
<td>0.28–0.29</td>
<td>0.14–0.17</td>
<td>0.27–0.29</td>
</tr>
<tr>
<td>Lopez-Mingues [54]</td>
<td>2017</td>
<td>Spain</td>
<td>56</td>
<td>50</td>
<td>52 (46–49)</td>
<td>Actigraphy</td>
<td>0.71</td>
<td>–0.12</td>
<td>0.65</td>
</tr>
<tr>
<td>Madrid-Valero [55]</td>
<td>2018</td>
<td>Spain</td>
<td>704</td>
<td>1446</td>
<td>53.7 (41–73)</td>
<td>Self-report</td>
<td>0.26</td>
<td>0.18</td>
<td>0.30</td>
</tr>
<tr>
<td>Markovic [33]</td>
<td>2018</td>
<td>Switzerland</td>
<td>36</td>
<td>22</td>
<td>13.2 ± 1.1</td>
<td>PSG</td>
<td>0.09</td>
<td>–0.28</td>
<td></td>
</tr>
<tr>
<td>Partinen [30]</td>
<td>1983</td>
<td>Finland</td>
<td>4476</td>
<td>9090</td>
<td>18+</td>
<td>Self-report</td>
<td>0.44</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>Setten [56]</td>
<td>2013</td>
<td>Spain</td>
<td>50</td>
<td>82</td>
<td>12</td>
<td>Actigraphy</td>
<td>0.64</td>
<td>0.38</td>
<td>0.65</td>
</tr>
<tr>
<td>Te Veleda [39]</td>
<td>2013</td>
<td>Netherlands</td>
<td>2372</td>
<td>1760</td>
<td>15.7 (12–20)</td>
<td>Self-report</td>
<td>0.54–0.56</td>
<td>0.25–0.28</td>
<td>0.34–0.36</td>
</tr>
<tr>
<td>Touchette [38]</td>
<td>2013</td>
<td>Canada</td>
<td>397</td>
<td>582</td>
<td>0.5–4</td>
<td>Parent-report</td>
<td>0.65–0.72</td>
<td>0.26–0.58</td>
<td>0.20–0.58</td>
</tr>
<tr>
<td>Watson [46]</td>
<td>2010</td>
<td>US</td>
<td>846</td>
<td>286</td>
<td>39.6 ± 15</td>
<td>Self-report</td>
<td>0.38</td>
<td>0.19</td>
<td>0.31</td>
</tr>
<tr>
<td>Webb [57]</td>
<td>1983</td>
<td>US</td>
<td>28</td>
<td>28</td>
<td>18.6–19.5</td>
<td>PSG</td>
<td>0.52</td>
<td>0.61</td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Characteristics of twin studies on sleep quality.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>nMZ</th>
<th>nDZ</th>
<th>Age, years (range)</th>
<th>Method</th>
<th>rMZ</th>
<th>rDZ</th>
<th>h²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barclay [15]</td>
<td>2010</td>
<td>UK</td>
<td>380</td>
<td>654</td>
<td>20 (18–27)</td>
<td>PSQI sleep quality component</td>
<td>0.49</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>UK</td>
<td>380</td>
<td>654</td>
<td>20 (18–27)</td>
<td>PSQI</td>
<td>0.42</td>
<td>0.25</td>
<td>0.41</td>
</tr>
<tr>
<td>Boomsma [35]</td>
<td>2015</td>
<td>Netherlands</td>
<td>218</td>
<td>256</td>
<td>Dutch Groningen Sleep Questionnaire</td>
<td>31</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaspers [49]</td>
<td>2017</td>
<td>US</td>
<td>102</td>
<td>98</td>
<td>(16–65)</td>
<td>PSQI</td>
<td>0.35</td>
<td>0.16</td>
<td>0.36</td>
</tr>
<tr>
<td>Genderson [17]</td>
<td>2013</td>
<td>Canada</td>
<td>604</td>
<td>534</td>
<td>(16–65)</td>
<td>PSQI</td>
<td>0.34</td>
<td>0.17</td>
<td>0.34</td>
</tr>
<tr>
<td>Gregory [48]</td>
<td>2017</td>
<td>UK</td>
<td>1226</td>
<td>1004</td>
<td>1-item categorical measure</td>
<td>17–88</td>
<td>0.31</td>
<td>0.25</td>
<td>0.32</td>
</tr>
<tr>
<td>Heath [41]</td>
<td>1990</td>
<td>Australia</td>
<td>1130</td>
<td>704</td>
<td>1-item categorical measure</td>
<td>18</td>
<td>0.34</td>
<td>0.13</td>
<td>0.33</td>
</tr>
<tr>
<td>Hu [31]</td>
<td>2019</td>
<td>US</td>
<td>164</td>
<td>130</td>
<td>(34–82)</td>
<td>PSQI &amp; PSQI sleep quality component</td>
<td>0.28</td>
<td>0.07</td>
<td>0.26</td>
</tr>
<tr>
<td>Madrid-Valero [55]</td>
<td>2018</td>
<td>Spain</td>
<td>704</td>
<td>1446</td>
<td>(41–73)</td>
<td>PSQI &amp; PSQI sleep quality component</td>
<td>0.35</td>
<td>0.11–0.133</td>
<td>0.31–0.34</td>
</tr>
<tr>
<td>Partinen [30]</td>
<td>1983</td>
<td>Finland</td>
<td>4476</td>
<td>9090</td>
<td>1-item categorical measure</td>
<td>&gt;18</td>
<td>0.47</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Pauw [18]</td>
<td>2009</td>
<td>Finlad</td>
<td>8628</td>
<td>2168</td>
<td>1-item categorical measure</td>
<td>33</td>
<td>0.33–0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor [32]</td>
<td>2015</td>
<td>US</td>
<td>3444</td>
<td>6156</td>
<td>PSQI</td>
<td>16</td>
<td>0.43</td>
<td>0.16</td>
<td>0.41</td>
</tr>
<tr>
<td>Ordonana [36]</td>
<td>2011</td>
<td>Spain</td>
<td>418</td>
<td>444</td>
<td>PSQI</td>
<td>PSQI</td>
<td>0.44</td>
<td>0.18</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Statistical analyses

Statistical analyses were run in R (version 3.5.1) with the ‘Metafor’ package using a random-effects model for heritability analyses, and a mixed-effects model to examine the influence of any moderators. The R script and datasets are available in the supplementary materials. To account for studies providing multiple effect estimates of the same cohort (e.g., longitudinal data, multiple assessment methods, subsamples of same cohort), a multi-level meta-analysis was performed which takes into account this dependency. This has been suggested to increase power and utilize maximum information in data [23].

Commonly reported effect estimates in behavioral genetic twin research are raw intraclass correlations in MZ and DZ twins, as well as the resulting proportions of variance attributed to A, C and E estimated from these correlations. Multiple studies only presented their best fitting model (dropping non-significant parameters), and reported only the variance decomposition based on this best fitting model. This model choice and preference is sensitive to sample size, thereby possibly presenting a biased perspective (often an over-estimation) of genetic influences on sleep [24]. We therefore present results of the meta-analysis of the twin correlations (MZ correlation and DZ correlation, respectively) as main results. In additional analyses, however, we also meta-analyzed the standardized variance components from both the full models and the best fitting models for papers that reported these statistics. Thus, we performed separate multi-level meta-analyses for rMZ and rDZ correlations, as well as heritability estimates from the full (A = h²_full) and the best fitting models (A = h²_best) where available. We transformed the raw rMZ, rDZ into Fisher’s Z scores which are assumed to be normally distributed – an assumption which is required to accurately derive estimates of mean effect sizes, and to ensure statistical tests are unbiased [25]. This method is preferred over conducting a meta-analysis directly on the correlations because the standard error of a
Results

Description of included studies

Of 5644 publications screened based on title and abstract, 5485 were excluded. The full-texts of the remaining 160 were independently read by the first two authors to assess eligibility. There were differences in the inclusion between the two authors for 24 papers on sleep duration and nine papers on sleep quality, after which an additional six were excluded. Reasons for exclusion are outlined in the PRISMA Flowchart (Fig. 1). In total 23 papers were included in the meta-analysis of sleep duration and 13 in the meta-analysis of sleep quality heritability. Descriptive information including twin correlations of the studies included in the meta-analyses are reported in Table 1 for sleep duration and Table 2 for sleep quality.

Sleep duration

Of the 23 papers included in the meta-analysis, 19 reported on independent studies. Five studies reported results from the USA, three from the UK, two from Australia, Italy and Finland, and one each from Switzerland, Croatia, China, Spain, The Netherlands and Canada. Papers were published between 1979 [27] and 2019 [28]. Sample sizes ranged between 26 twin pairs [29] and 6783 pairs [30]. The total sample size includes 45,328 individuals, including 9277 MZ twin pairs and 13,387 DZ twin pairs, aged 6 mo to 88 y. Sleep duration was measured by various measures (sleep duration component of the Pittsburgh Sleep Quality Index (PSQI), n = 3; self-reported average over a typical week, n = 1; self-reported average on a night, n = 2; self-reported categorical outcome, e.g., less than 4 h, 5, 6, or 7 h, or longer than 8 h, n = 2; sleep diary weekly average, n = 1; actigraphy, n = 5; parent-reported from one item of the Child Sleep Habits Questionnaire, Brief Infant Sleep Questionnaire or other infant sleep questionnaire, n = 4; Polysomnography (PSG), n = 2). Five studies provided MZ and DZ correlations separately for males and females, and one study included females only. All other studies combined estimates for males and females. Three of the studies reported the heritability of sleep duration in infancy; three in young childhood; four in middle childhood; one in older childhood; five in adolescence; nine in adulthood; and one in older adulthood.

Sleep quality

Thirteen papers from 10 independent samples fulfilled the inclusion criteria. From these studies 10 MZ and DZ correlations were used to estimate a pooled heritability estimate for sleep quality. Three samples were from the UK, three from the USA, and one each from Australia, Spain, The Netherlands and Finland. Papers were published between 1983 [30] and 2019 [31]. Sample sizes ranged between 147 twin pairs [31] to 6783 pairs [30]. The total sample size consists of 39,020 individuals, including 7648 MZ twin pairs and 11,862 DZ twin pairs, aged 16–88 y.

Sleep quality was also assessed by various measures. Seven studies used the PSQI global score as a measure of sleep quality, two papers used the Subjective Sleep Quality component score from the PSQI, and two papers used single questions with categorical responses. Seven studies pooled estimates for males and females, two studies reported both pooled and separate estimates for males and females, and one study included only females. Nine of the 10 papers examined the heritability of sleep quality in adulthood (ranging from 18 y to 88 y), and one examined adolescence (16 y) [32].

Meta-analysis of heritability estimates

Sleep duration

The 19 independent studies provided 65 MZ and 65 DZ correlations (Fig. 2), as well as 31 heritability estimates based on structural equation twin model-fitting approaches. MZ correlations ranged between 0.05 [16] and 0.97 [27]; DZ correlations ranged between −0.28 [33] and 0.92 [34]. Standardized heritability estimates (where available) ranged from 0% [15] to 71% [16].

The multi-level meta-analysis of sleep duration yielded an overall MZ correlation of 0.63 (Fisher’s Z-score rMZ = 0.75, SE = 0.12, t = 6.50, p < 0.001, 95% CI = 0.53–0.97) and an overall DZ correlation of 0.40 (Fisher’s Z-score rDZ = 0.43, SE = 0.10, t = 4.81, p < 0.001, 95% CI = 0.25–0.61). Based on these meta-analyzed correlations an overall heritability of sleep duration was estimated at 46%. This indicates that 46% of differences between individuals in sleep duration is due to differences in their genetic make-up. The magnitude of the difference between MZ and DZ correlations indicates very little contribution of the shared environment in sleep duration, and that non-shared environmental influences contribute the remaining 54% of variability in sleep duration. Heterogeneity was high (I^2 = 98%) for correlations in both monozygotic and dizygotic twins. Similar estimates (rMZ = 0.64, rDZ = 0.41, h^2 = 47.2%) were obtained in sensitivity analyses excluding overlapping samples and samples <30 participants. The meta-analysis of variance components yielded somewhat lower heritability estimates, namely 31.1% (95% CI = 0.21–0.41) based on the full model, and 33% (0.27–0.41) based on the best model.

Moderator analyses were then performed to explore sources of heterogeneity and determine whether sex, age or informant moderated the MZ and DZ correlations in sleep duration (see Table 3). These showed that MZ and DZ correlations were not significantly different for males and females, but did differ per age group and method of assessment. F(6, 58) = 3.37, p = 0.011 for rMZ moderated by age group; and F(6, 58) = 6.37, p < 0.001 for rDZ moderated by age group; F(4, 60) = 7.10, p < 0.001 for rDZ moderated by informant (there was no significant moderator effect of informant on rMZ). Heritability estimates for sleep duration were lowest in infancy and early childhood (17% and 20% respectively), increased to 41% and 52% in middle childhood, to 68% in adolescence, and again decreased in adulthood to 42–45%. Heritability estimates for sleep duration were also moderated by the method of assessment; such that heritability of parent reported sleep duration (8%) was much lower than heritability of self-reported sleep duration (38%) or sleep diary (52%). Heritability estimate of actigraphically measured sleep was highest (100%), whereas based on two relatively small studies heritability of sleep measured with PSG was 27% (see Table 3).

Sleep quality

The 13 papers provided 20 MZ and 20 DZ correlations (Fig. 3), as well as 16 heritability estimates based on structural equation twin
model-fitting approaches. MZ correlations ranged between 0.20 (for non-cohabiting male twins aged 25–y [30]) and 0.68 (for cohabiting male twins aged 25–y [30]). DZ correlations ranged between 0.06 (for non-cohabiting male twins aged 18–24 y [30]) and 0.27 (for cohabiting female twins aged 25–y [30]). Standardized broad-sense heritability estimates (where available) ranged from 0% [35] to 53% [18]. There was little contribution of the shared environment, which was estimated at 0% in two studies [15,35]. Non-shared environmental influences contributed between 56% [36] and 100% of variability in Sleep Quality symptoms [35].

The multi-level meta-analysis of sleep quality yielded an overall MZ correlation of 0.38 (Fisher’s Z-score rMZ = 0.40, SE = 0.02, t = 16.92, p < 0.001, 95% CI = 0.35 – 0.44) and an overall DZ correlation of 0.16 (Fisher’s Z-score rDZ = 0.16, SE = 0.02, t = 10.38, p < 0.001, 95% CI = 0.13 – 0.19). Based on these meta-analyzed correlations an overall heritability of sleep quality was estimated at 44%. This indicates that 44% of differences between individuals in Sleep Quality is due to differences in their genetic make-up. The magnitude of the difference between MZ and DZ correlations indicates that the shared environment has only a small contribution for the individual variability in sleep quality, and that non-shared environmental influences contribute the remaining 56% of variability in sleep quality. Heterogeneity was relatively high for correlations in both monozygotic (I² = 60%) and dizygotic (75%) twins. Similar estimates (rMZ = 0.37, rDZ = 0.16, h² = 41.8%) were obtained in sensitivity analyses excluding overlapping samples and samples <30 participants. The meta-analysis of variance components yielded lower heritability estimates for sleep quality, namely 30.6% (95% CI = 0.12 – 0.48) based on the full model, and 33.3% (95% CI = 0.26 – 0.41) based on the best model.
Moderator analyses were then performed to explore sources of heterogeneity and to determine whether sex or age moderated the MZ and DZ correlations in sleep quality (see Table 4). These showed that sex and age did not moderate the MZ correlations nor DZ correlations. Hence, MZ and DZ correlations in these studies are not significantly different for males and females, nor did these differ per age group.

Discussion

Summary of main findings

In the present meta-analyses, we synthesized research on the heritability of sleep duration and quality utilizing monozygotic and dizygotic twin correlations from 28 independent twin samples (19 including sleep duration and 10 including sleep quality), covering an age range from 6 mo to 88 y for sleep duration, and 16 y—95 y for sleep quality, with a total sample size of 45,328 and 39,020 individual twins for sleep duration and quality, respectively.

The meta-analysis indicated that the 46% of the variability in sleep duration can be attributed to genetic variation between individuals, and this estimate is moderated by age group, and method of assessment, but is similar across the sexes. Heritability estimates were lowest in infancy and young childhood, highest in adolescence and decreased again in adulthood. Sleep duration heritability estimates were highest for actigraphically measured sleep duration, moderate for self-reported and diary reported sleep duration, and lowest for parent-reported and PSG measured sleep duration (though the latter estimate was based on two small samples). Heritability of subjective sleep quality was estimated to be 44%, and this estimate did not differ across age and sex, and all samples measured sleep quality by self-report. The pooled heritability estimates based on fitted twin models were lower both for sleep duration (33%) and sleep quality (33%).

Interpretation of findings

Our findings indicate that sleep characteristics are under moderate to strong genetic control. Specifically, monozygotic twin correlations compared to dizygotic twin correlations were 1.7 and 2.4 times as large for sleep duration and sleep quality, respectively. This is consistent with the findings of a recent independent meta-analysis with similar methodology [19]. This provides evidence for the importance of additive genetic factors for sleep duration, and the possibility of additive and non-additive genetic factors in sleep quality. Additionally, the magnitude of difference between the twin correlations suggests that the unique environment of the twin individual plays a notably larger role than the environment the twins share, such as parenting practices. A comprehensive examination of environmental factors (e.g., stressful life events) that may impact sleep should be the focus of future research. That said, the contribution of the shared environment for the variability in sleep duration was generally higher in pediatric cohorts. For example, the proportion of variance in parent-reported nighttime sleep duration attributed to the shared environment was estimated to be 81% at 12 mo [34], 66% at 15 mo [37] and 48% at 18 mo [38]. Te Velde et al. [39] also reported, albeit with cross-sectional data, that the shared environment was only important for sleep duration during early adolescence, but of negligible importance from age 16 onwards. In addition, both studies reporting longitudinal changes in heritability of sleep duration during childhood, confirmed that the role of the shared environment indeed decreases with age (from 81% at 12 mo [40] to 60% at 5 y [34]; and from 48% at 18 mo to 17% at 4 y [38]). The role of the shared environment was consistently negligible for sleep quality, although this could be because of the lack of behavioral genetic studies evaluating the heritability of sleep quality in pediatric population, which in turn is a result of the lack of a valid measure thereof. These findings indicate that sleep interventions aimed at optimizing sleep duration should ideally be early and family-based. Whether this is the same for sleep quality remains unclear until this phenotype is assessed in pediatric populations using behavioral genetic approaches.

Heritability in both sleep characteristics were not moderated by sex. This is in line with the previous meta-analysis of twin-based heritability [19], and with previous GWAS studies of sleep duration [3] and insomnia [8], showing a high genetic correlation between the sexes. However, sex-differences in the heritability of sleep quality were found in the Finnish Twin Cohort [18,30], though these were not confirmed in other studies, despite stratifying analyses by sex [41,42]. It is important to note however, that these findings do not indicate that sleep duration and quality do not differ between males and females, but that the extent to which their etiology is due to genetic factors is similar for males and females. Whereas women report longer sleep duration and worse sleep quality across studies and cultures, heritability estimates examine relative differences in variance rather than mean differences.

![Forest plots of the monozygotic and dizygotic twin pair correlations for sleep duration from 65 effects with standard error bars.](image-url)
between the sexes. These findings thus indicate that the genetic etiologies of both sleep duration and quality is shared between males and females.

Furthermore, the heritability of sleep duration differed with age and with method of assessment, which contradicts the findings of the recent meta-analysis of heritability of sleep duration [19], where age and method of assessment were not significant moderators. This is likely due to the fact that in this previous meta-analysis, studies in pediatric cohorts (<6 y) were not included. An increase in the heritability of sleep duration was also

<table>
<thead>
<tr>
<th>Moderator</th>
<th>Categories</th>
<th>MZ</th>
<th>DZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Male</td>
<td>0.42</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.48</td>
<td>0.45</td>
</tr>
<tr>
<td>Age</td>
<td>Infancy</td>
<td>1.27</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Young childhood</td>
<td>1.14</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Middle childhood</td>
<td>0.96</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Older childhood</td>
<td>0.76</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Adolescence</td>
<td>1.09</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Adulthood</td>
<td>0.40</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Older adulthood</td>
<td>0.36</td>
<td>0.35</td>
</tr>
<tr>
<td>Informant</td>
<td>Parent-reported</td>
<td>0.85</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Self-reported</td>
<td>0.79</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Actigraphy</td>
<td>0.79</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Sleep diary</td>
<td>0.31</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>PSG</td>
<td>0.23</td>
<td>0.23</td>
</tr>
</tbody>
</table>

DZ = dizygotic twins; h² = heritability estimate calculated as 2(rMZ-rDZ); MZ = monozygotic twins; PSG = Polysomnography rDZ = dizygotic twin correlation; rMZ = monozygotic twin correlation; 95% CI = 95% confidence interval.
reported in longitudinal analyses in pediatric cohorts [38,40], and mirrored in our moderator analyses. Notably, this variability in heritability with age has also been observed for traits such as temperament, Intelligence Quotient (IQ), and weight (see Plomin et al. [43]). In the case of sleep duration however, this seems to reach a peak during adolescence, and the unique environment an individual experience seems to play a crucial role thereafter. More importantly, the differences in heritability estimates across methods of assessment could potentially impact future sleep research and inferences drawn from it. These findings could indicate that estimates of sleep duration assessed via different methods (i.e., self-report vs. actigraphy, vs. PSG) can represent different physiological phenomena that are under different levels of genetic control. Though PSG is the gold standard for estimating sleep, based on twin research conducted thus far, it seems that PSG-estimated sleep duration is under weaker genetic control (27%). With polysomnography, however, there is also a well-established first night effect where individuals sleep worse than usual during the first night in the sleep lab. This reactivity could also inflate environmental influences in twin studies leading to artificially low heritability estimates. Self-reported measures are less prone to such reactivity bias. Of note, however, one of the two polysomnographic studies did perform an adaptation night in the sleep lab [42]. In addition, it could be that the differences in heritability estimates could reflect measurement error. For example, the remarkably low heritability estimate of parent-reported sleep duration may reflect inaccuracy in parental reports of children’s sleep. We report substantial differences in heritability of sleep duration across different methods, i.e., around half of the variability in self-reported sleep duration and all of the variability in actigraphically measured sleep duration is genetically influenced. Single Nucleotide Polymorphism (SNP)-based heritability for sleep duration has also been estimated recently, ranging from 9.8% for self-reported sleep duration [5] to 19% for sleep duration estimated with actigraphy [6]. This could indicate that inactivity, the basis of actigraphic sleep duration estimates, is under stronger genetic control than self-reported sleep duration. Our unusually high heritability estimate of actigraphically assessed sleep duration (100%) arises from a couple of the studies included in that reported negative DZ twin correlations (sometimes interpreted as genetic factors that contribute to differences in a trait). These inflate the difference between rMZ and rDZ leading to a heritability estimate ≥100%. All studies using actigraphy and PSG also had small sample sizes (ranging from 26 pairs to 206 pairs), calling for larger twin studies of objectively measured sleep duration. Though it has been estimated that it would require 50 million people to explain >90% of the heritability [44], as sample sizes are increasing exponentially this number may be reached in the next decade. Self-report, however, remains a core tool for assessing sleep both in research and clinical practice, thus it remains crucial to understand the extent to which self-reported traits are under genetic control and the underlying genetic architecture.

Genetic overlap between sleep characteristics and other traits

Many studies included in the meta-analyses evaluated genetic overlap between sleep duration or quality and other traits. Though these effects could not be formally analyzed, we will briefly summarize them here. Butkovic and colleagues [45] showed that sleep duration and personality traits have common genetic influences. Specifically, in this study adolescent twins that sleep shorter had higher neuroticism and openness scores that those sleeping longer (i.e., above 6.5 h), and these phenotypic associations were mainly mediated by genetically overlapping factors. In addition, Watson and colleagues [46], showed in 1224 adult males that short sleep is associated with higher Body Mass Index (BMI) after careful adjustment for the influence of genetic factors and the shared environment. This association was confirmed in Chinese women [47], with more specific measures of adiposity (assessed with Dual-energy X-ray absorptiometry (DXA) scans).

Furthermore, studies included in our meta-analysis showed that subjective sleep quality shares common genetic underpinnings with several psychiatric traits. Gregory and colleagues [48] showed that the longitudinal association between Attention Deficit Hyperactivity Disorder (ADHD) symptoms and poor sleep quality between 5 and 18 y, were due to genetic influences (55%) and non-shared environmental (45%) influences. Similarly, according to a study in 400 adult twins [49] the phenotypic correlation between sleep and depression, can be attributed to genetic effects (60%), and non-shared environmental influences (40%). This finding is mirrored in a study by Gregory and colleagues [16] which demonstrated between that 50%—90% of the associations between insomnia and depression is accounted for by shared genes [30]. Notably, Taylor and colleagues [32] also showed that both genetic and environmental influences of psychotic symptoms (e.g., paranoia, hallucinations and cognitive disorganization) displayed a moderate degree of overlap with those of sleep quality.

Limitations

Some limitations of our study must be noted. First, heritability estimates depend on the variance of the population studied [13]. Given that the included studies were conducted in developed countries from Europe, USA and Australia, these estimates might not be generalizable to other populations. Second, measurement error in the assessment of sleep duration and quality might influence the heritability estimates. Indeed, we showed that heritability estimates for sleep duration differed per method of assessment, which could in part be due to measurement error. Third, two studies which met our inclusion criteria did not report twin correlations, only variance components from the best-fitting model (and hence were not included in our meta-analysis), possibly generating selection bias in the overall heritability estimates. Therefore, we also meta-analyzed heritability estimates based on structural equation twin model-fitting approaches, which yielded smaller heritability estimates. Nevertheless, a previous large-scale meta-analysis of heritability [51] showed that estimates of variance components from model-fitting can underestimate the true trait heritability, when compared with heritability based on twin correlations, thus these results should be interpreted with caution.

Conclusions

Sleep duration and quality both have a strong genetic component, explaining 46% and 44% of the variation across individuals, respectively. The remaining variation in these sleep characteristics can be attributed to non-shared environmental influences. The shared environment seems to play a role for individual variation in sleep duration in early childhood only. While sleep characteristics may share genetic underpinnings with other markers of physical and psychological health, these should be further studied in order to inform directionality of phenotypic associations. Our findings highlight the importance of considering genetic influences when aiming to understand the underlying mechanisms contributing to the development of sleep patterns across the lifespan.
Practice points

If the high genetic control of sleep characteristics is confirmed, recommendations for appropriate sleep duration should be directed towards other sleep characteristics (e.g., sleep quality) that also voluntary on the environment, and perhaps those that are under voluntary control (such as the timing of bedtime).

Research agenda

Our research highlights several gaps that future research within and beyond behavioral genetics should focus on. First, more studies should quantify the heritability of more objective estimates of sleep duration (actigraphy and polysomnography). Second, efforts should be made to define sleep quality in pediatric populations, and its heritability should be subsequently quantified. Finally, more intervention studies should focus on the early ages, as this is a period when the effect of the environment is more pronounced.

Conflicts of interest

The authors have no conflicts of interest to disclose.

Disclosure

The work was funded by NWA Startimpuls Royal Netherlands Academy of Arts and Sciences 2017 Grant (AZ/3137) and by the European Research Council grant ERC-2014-AdG-671084 INSOMNIA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.smrv.2021.101448.

References


Madrid-Valero JJ, Sanchez-Romera JF, Gregory AM, Martinez-Selva JM, Ordonana JR. Heritability of sleep quality in a middle-aged twin sample from Spain. Sleep 2018;41(9).

