Neuroscience and Biobehavioral Reviews 109 (2020) 78–89

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

# Twin studies of subjective sleep quality and sleep duration, and their behavioral correlates: Systematic review and meta-analysis of heritability estimates

Juan J Madrid-Valero[a,b], María Rubio-Aparicio[c], Alice M. Gregory[d], Julio Sánchez-Meca[e], Juan R. Ordoñana[a,b,*]

[a] Department of Human Anatomy and Psychobiology, University of Murcia, Spain
[b] Murcia Institute of Biomedical Research, IMIB-Arrixaca, Spain
[c] Department of Health Psychology, Faculty of Health Science, University of Alicante, Spain
[d] Department of Psychology, Goldsmiths, University of London, London, United Kingdom
[e] Department of Basic Psychology and Methodology, University of Murcia, Spain

ARTICLE INFO

*Keywords:*
Heritability
Meta-Analysis
Twin
Sleep quality
Sleep duration

ABSTRACT

Twin studies have shown that a substantial proportion of the variance for sleep variables is due to genetic factors. However, there is still considerable heterogeneity among research reports. Our main objectives were to: 1) Review the twin literature regarding sleep quality and duration, as well as their behavioural correlates; 2) Estimate the mean heritability of subjective sleep quality and sleep duration; 3) Assess heterogeneity among studies on these topics; and 4) Search for moderator variables. Two parallel meta-analyses were carried out for sleep quality and sleep duration. Seventeen articles were included in the meta-analysis. Mean MZ correlations were consistently higher than DZ correlations. A mean heritability of 0.31 (95% CI: 0.20, 0.41) was found for subjective sleep quality (range: 0-0.43) and 0.38 (95% CI: 0.16, 0.56) for sleep duration (range: 0-1). Heterogeneity indexes were significant for both sleep quality ($I^2$ = 98.77, $p < .001$) and sleep duration ($I^2$ = 99.73, $p < .001$). The high heterogeneity warrants further research considering possible moderators that may affect heritability.

## 1. Introduction

Sleep, or some kind of rest in different forms, can be found universally amongst different animal species (Cirelli & Tononi, 2008). It is important for the optimal functioning of an organism and exerts its effects through the modification of brain activity. Overall, sleep appears to be essential for existence. However, we remain far from understanding the exact purpose and the specific functions fulfilled by sleep. Given its alleged universality, it is thought that sleep evolved early and that its role must be the same, or very similar, in every animal. In spite of this, there is considerable variability in sleep both inter and intra-species (Cirelli & Tononi, 2008). Indeed, humans differ in many features of sleep. Overall sleep quality, duration, depth, efficiency, restorative value, amongst other factors, present substantial inter-individual variation (Barclay et al., 2010b; Madrid-Valero et al., 2018; Ohayon et al., 2004). Differences in age, sex, bedtime and wake-time habits, lifestyle choices, job schedule, partner, or general health are amongst factors that can contribute to such variation (Ohayon et al., 2004). For example, sleep characteristics change dramatically at different stages of life for humans. Children up to preschool typically require more than 10 hours of sleep whereas older adults typically need around 7-8 hours (Hirshkowitz et al., 2015). Similarly, there are differences in sleep architecture at different stages of the life-course (Ohayon et al., 2004), and sleep duration and quality typically decrease from adolescence to adulthood while sleep disturbances tend to increase reaching a maximum in older adults (Grandner, 2019; Madrid-Valero et al., 2017a). Sex is another variable that has been linked to the way that we sleep, although sex-related variations are not consistent across measures (Meers et al., 2019). Overall sex differences in short sleep duration are small and inconsistent (Grandner, 2019). Healthy females usually show better results than males when sleep parameters (e.g., sleep quality or duration) are measured objectively (Mong and Cusmano, 2016). Paradoxically, women often report poorer subjective sleep quality than men and a higher prevalence of self-reported sleep

* Corresponding author at: Faculty of Psychology, The University of Murcia, Campus de Espinardo, 30100, Murcia, Spain.
*E-mail address:* ordonana@um.es (J.R. Ordoñana).

https://doi.org/10.1016/j.neubiorev.2019.12.028
Received 27 July 2019; Received in revised form 4 December 2019; Accepted 23 December 2019
Available online 30 December 2019

disturbances (Madrid-Valero et al., 2017a; Meers et al., 2019; Schredl & Reinhard, 2011; Zhang & Wing, 2006). There are also a wide variety of environmental differences that influence different aspects of sleep. These involve cultural habits including differences in TV prime time, commercial opening hours, time zone, and latitude. For example, bed and wake time vary widely across countries (Walch et al., 2016) and some behaviors such as siesta are common in some cultures (e.g. Mediterranean countries) but not in others (Sayón-Orea et al., 2013).

Additionally, given the evolutionary character and universality of sleep, genetic factors are thought to play a major role in sleep-related phenotypes. In fact, the search for understanding the role of genetic factors related to sleep has been advancing steadily during the previous decades thanks to the expansion of molecular and quantitative genetic studies. The former has developed new ways of specifying genetic variants associated with sleep phenotypes; while the latter continues to disentangling the relative role of genetic and environmental factors in phenotypic variability. Altogether, both approaches have resulted in important new knowledge about the genetics of sleep.

Towards this endeavor, twin studies have proven an invaluable resource and their number and scientific utility has grown alongside the development of new and existing twin registries (Odintsova et al., 2018). The main objective of a classical twin study is to disentangle the role of genetic and environmental factors in the development of individual differences (Knopik et al., 2017). Hence, a number of twin studies have analysed the heritability of sleep-related phenotypes (Barclay & Gregory, 2013; Polderman et al., 2015). These have provided information about the relative weight of genetic and environmental factors for explaining individual differences in these features of sleep. Furthermore, twin data can be used to address scientific questions that go beyond estimating heritability for a single variable. Causal inference, multivariate genetic correlations, or gene-environment interplay can be analysed using this kind of genetically informative designs. Additionally, twins are sometimes participants of molecular genetic studies and, in particular, epigenetic studies.

Virtually every aspect of sleep, including those assessed using both objective and subjective measures, has been analyzed in order to advance knowledge of genetic factors. These include the magnitude of genetic influences as well as the way genes exert their influence. Furthermore, research has suggested moderate and significant heritability for chronobiological variables, such as morningness-eveningness preference or circadian rhythmicity (López-Mínguez et al., 2016; López-Mínguez et al., 2015; Vink et al., 2001); sleep neurophysiology (Rusterholz et al., 2018); sleep disorders or dysfunctions (Desai et al., 2004; Lessov-Schlaggar et al., 2008; Lind et al., 2015); and particular cultural habits, such as siesta (López-Mínguez et al., 2017). Of course, there are also important variations in such estimates. Since heritability refers to the proportion of the variance of a trait explained by genetic factors in a specific population at a specific moment (Visscher et al., 2008) it is unsurprising that it varies across different populations.

Among sleep measures, the variability in sleep quality and sleep duration has been assessed particularly widely in general population/non-clinical samples. However, the apparent agreement about their importance as sleep phenotypes is not paralleled by a consensus about how to define and measure them. Definitions have not yet been harmonized, especially for sleep quality, and researchers have used different techniques for their appraisal: from self-report, using a single question or a more elaborated questionnaire, to inferences from objective measurements (Hirshkowitz et al., 2015; Krystal & Edinger, 2008; Lauderdale et al., 2008; Ohayon et al., 2017). Actually, both sleep quality and length encompass several aspects of sleep. For example, measures of sleep quality typically comprise information about sleep duration, latency, and efficiency, among other features of sleep (Grandner, 2019; Krystal & Edinger, 2008). Similarly, sleep duration reflects different aspects of sleep, including chronobiological (e.g., patterns of rhythmicity), environmental (e.g., social pressures for bedtime and wake-up timing) and functional (e.g., melatonin levels or sleep phase) components.

Reflecting this lack of specificity, the subjective and objective components of the sleep experience, as related to sleep quality and duration, do not correlate well. This has been interpreted as subjective and objective measures capturing different dimensions of the sleep experience (Aili et al., 2017; Buysse et al., 2008; Jackowska et al., 2011; Tubs et al., 2019). Regardless of the measurement method, sleep quality and duration appear to be closely related. As discussed, duration is usually included among the dimensions comprising sleep quality. Consequently, these variables are phenotypically, and also genetically correlated (Barclay et al., 2010b).

Despite the lack of a rigorous definition and the diversity of measurements, both sleep quality and duration have been analyzed from a genetically informed perspective and twin studies have helped to determine the relative weight of genetic and environmental factors on their variability and the associations with other behavioral correlates.

### 1.1. Heritability of sleep quality and sleep duration

The heritability of sleep quality and sleep duration has been estimated in different populations and at different stages of life. Most studies have reported a moderate heritability for sleep quality. For example, genetic factors have been found to account for 41% of the variance in sleep quality in adolescents (Taylor, Gregory et al., 2015), 43% in young adults (Barclay et al., 2010b), and 34% in adults (Genderson et al., 2013; Madrid-Valero et al., 2018). However, not all studies reach this conclusion and in one adult twin study no genetic factors were found (Boomsma, van Someren et al., 2008). The authors suggest that these results could be due to the use of a single item to measure night time behavior that might result in substantial measurement error. Instead, in this study, all the variance was accounted for by non-shared environmental influences, which are those which make members of a family less alike (these estimates include measurement error).

When it comes to sleep duration, it is more difficult to draw out trends. A heritability of 65% was found in 12 years old twins (Sletten et al., 2013), 63% in adolescents/early adult twins (Butkovic et al., 2014), and 26-29% in adults (Genderson et al., 2013; Madrid-Valero et al., 2018). However, as with sleep quality, certain studies did not find any significant contribution of genetic factors to sleep duration (Barclay et al., 2010b; Boomsma et al., 2008).

Twin studies using actigraphy to measure sleep parameters have reached similar estimates. López-Mínguez et al. (2017), using a composite measure (considering wrist temperature, activity and position), found a heritability of 65% and 61% for duration of night-time and diurnal sleep respectively. More recently, another study using actigraphy reported 49% heritability for total sleep and 44% for the mean duration of sleep episodes (Gehrman et al., 2019).

Genome-wide association studies have also addressed this question. These studies have found SNP-based heritabilities (i.e., the proportion of phenotypic variance explained by all single nucleotide polymorphisms – SNPs - assessed using genotyping arrays) of around 10% for sleep duration and sleepiness and ranging from 7 to 20% for insomnia (Dashti et al., 2019; Jansen et al., 2019; Lane et al., 2017). These estimations are notably lower than those from twin studies, as has been found for most other phenotypes that have been assessed in this way. This represents an example of the, so-called, “missing heritability” issue (Manolio et al., 2009; Mayhew & Meyre, 2017). Explanations given for the discrepancies in heritability estimates between twin and SNP-based heritability studies, include that SNP methods fail to capture a large number of common variants with small effect as well as rare variants with large effect. Alternatively, heritability estimated in twin and family studies may include specific effects, such as gene-environment correlation, not captured by SNP-based methods (Yang et al., 2017).

### 1.2. Externalizing behaviors and sleep

Sleep problems have been linked with a wide variety of externalizing behaviors including ADHD (Attention Deficit Hyperactivity Disorder), aggression, conduct disorders and addictions among other adverse consequences (Gregory & Sadeh, 2012, 2016).

Genetic and environmental influences on the relationship between ADHD and sleep quality have also been tested in adults. Gregory et al. (2017) found that children with ADHD have poorer sleep quality than those without when they reach adulthood - as long as ADHD persisted over time. The genetic and environmental influences on this relationship in adulthood were studied, finding that there is a substantial genetic correlation between ADHD and poor sleep quality (rA = 0.49) and that the phenotypic correlation between these two variables is almost equally explained by genetic and non-shared environmental factors.

Additionally, the relationship between sleep quality, diurnal preference and externalizing behaviors was tested in a sample of young adult twins and siblings. This study found that there was an association between poor sleep quality and externalizing behaviors. Likewise, there was an association between greater eveningness and externalizing behaviors. For both sleep quality or morningness-eveningness, the relationship with externalizing behaviors seemed to be mainly explained by genetic factors (around 80%) (Barclay et al., 2011).

In another twin study of children aged between 6 and 12 years, the relationship between different measures of sleep and two dimensions of antisocial behavior was studied (Madrid-Valero et al., 2020). A strong association between sleep measures and both dimensions of antisocial behavior was found. However, the origin of these relationships was different. The genetic overlap between aggression and sleep variables was larger than the genetic overlap between sleep variables and rule breaking.

There is a shortage of studies studying sleep and externalizing behaviors in middle-aged populations. However, a couple of studies found that there is some genetic overlap between externalizing behaviors (e.g. alcohol abuse and dependence and antisocial personality disorder) and insomnia (Lind et al., 2017) or diurnal preference (Watson et al., 2013).

### 1.3. Internalizing behaviors and sleep

The relationship between sleep and internalizing behaviors (e.g. depression and anxiety) has been widely studied using a behavioral genetic approach (Barclay & Gregory, 2013).

There is a robust phenotypic association between sleep quality and depression (Alvaro et al., 2013). Several studies have addressed genetic and environmental influences on this association. For example, in a sample of young adults, Gregory et al. (2011) found a high genetic correlation between poor sleep quality and symptoms of depression (rA = 0.68). Overall, the phenotypic association was mainly explained by genetic factors (58%). Similar results were found using a sample of middle-aged twins (Gasperi, Herbert et al., 2017).

In addition, the association between other sleep measures/ disorders (e.g. insomnia), and depression have also been considered from a genetically informed perspective. Studies focusing on the relationship between insomnia and depression have found that there is a substantial (or complete) genetic overlap between these two phenotypes (Gehrman et al., 2011; Lind et al., 2017). Moreover, Gregory et al. (2016) studied this relationship longitudinally in young adult twins and found that genetic factors were important in explaining the persistence of the association between these two traits (with genetic correlations above 0.73).

The relationship between chronotype and depression has also been considered in twin studies. Toomey et al. (2015) reported that genetic factors accounted for 59% of the significant phenotypic correlation (-0.15) where greater eveningness was associated with more depression symptoms. In a sample of elderly twins, a substantial genetic overlap (rA = 0.40) between depression and sleepiness was reported. This genetic correlation decreased (rA = 0.21) when the model took into account covariates including age, BMI and sleep apnea, among others (Lessov-Schlaggar et al., 2008). The relationship between sleep quality and emotional regulation was also tested in a co-twin design and a significant relationship between them was found, which seems to be influenced by common genetic factors (Medda et al., 2019). Finally, Watson et al. (2014), found that there is an interaction between sleep duration and symptoms of depression such that higher heritability was found for those subjects with short or long sleep duration.

A similar picture can be found for anxiety and sleep. There is also a robust relationship between these two traits (Alvaro et al., 2013) and both phenotypes are modestly heritable (Gregory et al., 2011; Madrid-Valero et al., 2018). In a sample of young adult twins, Gregory et al. (2011) found a moderate genetic correlation between sleep quality and symptoms of anxiety (rA = 0.58). Their phenotypic association was explained mainly by genetic factors (74%). Regarding insomnia, a complete genetic overlap between this sleep disorder and generalized anxiety disorder in a sample of adult twins has been reported (Lind et al., 2017). Additionally, in a sample of twins aged between 8 and 16 a complete genetic overlap between insomnia and overanxious disorder was found (Gehrman et al., 2011).

### 1.4. Sleep and other aspects of health

Sleep is related to almost all psychological disorders and sleep phenotypes (e.g. insomnia or hypersomnia) are included as diagnostic criterion for some (American Psychiatric Association, 2013; Plante, 2019). Many of the twin studies focusing on the associations between sleep and psychological disorders focus on mood (i.e. depression) and anxiety disorders, as reviewed previously. In contrast, there are far fewer studies exploring the genetic associations between sleep phenotypes and other mental health problems – although there are some reports of this type (for a discussion of sleep and externalising difficulties, see previously). There are also other examples of mental health symptoms associated with disturbed sleep, and Taylor et al. (2015) for example, reported a substantial genetic overlap between psychotic-like experiences and both poor sleep quality and insomnia. Similarly, recent GWAS (Genome-Wide Association Study) and PRSs (Polygenic Risk Score) analyses have reported a consistent link between sleep duration and psychiatric traits, most notably schizophrenia and bipolar disorders (Dashti et al., 2019; Jansen et al., 2019; Lane et al., 2017; Lewis et al., 2019). Butkovic et al. (2014) found that there is a modest genetic correlation between sleep duration and the personality traits neuroticism (rA= -0.26) and openness (rA= -0.22). However, there is a shortage of behavioral genetic studies addressing the relationship between sleep phenotypes and other aspects of psychopathology including eating disorders, personality disorders and substance use.

Some studies have also used twins to investigate the relationship between sleep and cognitive performance/deterioration, although they do not provide heritability estimates. They have generally found longitudinal associations between sleep duration (short or extended) and subsequent lower cognitive scores (Virta et al., 2013) or incidence of dementia (Bokenberger et al., 2017). Paunio et al. (2009) utilized a co-twin design and reported that poor sleep predicted life dissatisfaction assessed 6-years later, but not vice versa.

In addition to the strong links between sleep and mental health and function, sleep is also essential for our physical health (Luyster et al., 2012) – and this association has been addressed using twin samples. Both, sleep quality and duration have been associated with mortality risk. Hublin et al. (2007) reported, in different studies with the same sample, an increased risk of mortality for subjects in the upper and lower extremes of the sleep duration distribution and also in participants with poorer sleep quality (Hublin et al., 2011). Nonetheless, the scientific literature shows mixed results regarding mortality for subjects with long/short sleep duration or insomnia (Cappuccio et al., 2010; Liu

et al., 2017; Lovato & Lack, 2019). One explanation for these associations, is the possibility that there is mediation by other illnesses and disorders that have not been explored within the research studies. In addition to other sleep variables, respiratory (i.e. obstructive sleep apnoea) and vascular (i.e. restless leg syndrome) sleep problems have also been explored using twin studies, and they report heritabilities above .5 for these disorders (Desai et al., 2004). In a study from our research group, we found that there is a substantial genetic overlap between symptoms of obstructive sleep apnoea and depression, anxiety and externalising behaviors (Madrid-Valero et al., 2020).

Certain other variables are strongly associated with, or manifestations of, a wealth of health problems. These include body mass index (BMI) and pain, and both of them have been associated with sleep quality and duration in twin studies. In one study, a significant correlation was found between sleep duration and BMI (Watson, Buchwald et al., 2010). In a subsequent report from the same sample, the authors found that, besides a significant relationship between short sleep duration and increased BMI, the genetic influences on BMI were higher for those subjects with short sleep duration as compared to longer sleep duration. In other words, short sleep duration could increase the expression of genetic risk for high BMI (Watson et al., 2012). In addition to sleep duration, BMI has also been associated with sleep quality. The direction of effects (i.e. whether poor sleep quality leads to a BMI increase or whether a high BMI leads to sleep problems) has been explored using a co-twin design where the association between variables can be analyzed while controlling for the confounding effects of family factors. The results confirmed the significant relationship between sleep quality and BMI, even after applying high levels of control, including genetic factors. Moreover, a possible directionality of this relationship was suggested, such that sleep quality appeared to strongly affect BMI, while the opposite association appeared to be less robust and consistent in a nonclinical sample (Madrid-Valero et al., 2017b).

Sleep quality has also been related to pain. Moderate but significant phenotypic correlations (0.23 - 0.36) between pain and poor sleep quality, together with substantial genetic overlap between them (genetic correlations ranged from 0.33 to 0.69) have been found in samples of adult twins (Gasperi et al., 2017; Pinheiro et al., 2018).

### 1.5. Scope and objectives of the study

As noted, twin studies have produced a considerable amount of information that has been instrumental in our understanding of the genetic underpinnings of the variability in subjective sleep quality and sleep duration. The basic outcome of those studies has been the estimation of a moderate but significant heritability for both phenotypes. Nonetheless, there is considerable heterogeneity between estimates, so it is essential to synthesize available evidence on these research questions. Meta-analysis allows us to address such tasks (Rubio-Aparicio et al., 2018). Consequently, we have carried out two parallel systematic reviews and meta-analyses in order to determine how much of the variance in subjective sleep quality and duration is explained by genetic factors. Following the classic objectives of a meta-analysis (Borenstein et al., 2009) our main aims were to: 1) Review the contribution of twin research to scientific knowledge regarding sleep quality and sleep duration, as well as their behavioral correlates; 2) Estimate the mean heritability of subjective sleep quality and duration; 3) Assess heterogeneity among twin studies on these topics; and 4) Search for characteristics of the studies (moderator variables) that may be explaining heterogeneity.

## 2. Method

This meta-analysis has been reported following the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) statement (Moher et al., 2009). Supplementary Table 1 presents the PRISMA checklist for this meta-analysis.

### 2.1. Search strategy

The search was conducted from the 5th to 29th of November 2018 in PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) and Web of Science (https://apps.webofknowledge.com/). The following keywords were used: *sleep* AND *monozygotic/dizygotic/twin*/heritability*. The search was conducted without any filter for dates or language.

#### 2.1.1. Inclusion and exclusion criteria

Studies that provided an estimation of the heritability of subjective sleep quality or sleep duration published up until November 2018 were included. As there is a lack of consensus concerning the definition and measurement of sleep quality and duration, we focused on all self-report studies regardless of whether they included a single question or a standardized questionnaire to measure these variables. Studies inferring poor sleep quality by focusing on sleep problems/difficulties (e.g., utilizing clinical samples) were excluded. Studies utilizing objective measures to draw inferences about sleep quality were excluded but studies utilizing objective measures to draw conclusions about sleep length were included. In the case of sleep duration we included papers providing a measure of length of sleep during night-time, regardless of the specific measure and whether it was prospective or retrospective (see Table 1 for a summary of measures).

Regarding the exclusion criteria, studies that did not use a twin sample (e.g. GWAS studies) were excluded. Studies which used a sample with a mean age below 6 years old were also excluded since sleep characteristics (e.g. napping) and sleep duration in early stages of the life course are very different from those found in older populations (Hirshkowitz et al., 2015). We decided to focus on sleep in participants who were school-aged (an age at which routines are often well-established) and older. Only studies using independent samples were used. Therefore, duplicated samples were removed. Where more than one paper from a sample included data about sleep, we selected one to include in the meta-analysis based on sample size (we selected the larger sample size) and the details reported (e.g. whether the paper reported heritabilities for men and women separately – we selected the paper with greater detail). Finally, only papers containing relevant information regarding the heritability of sleep quality or sleep duration were included. Conference abstracts were assessed but eventually excluded since limited information was available and these conference papers were typically published as articles as well.

The search was conducted simultaneously for sleep quality and sleep duration. This search yielded 650 results in PubMed and 1109 in Web of Science. After removing duplicates, a total of 1275 records were screened. Of those, 101 articles were assessed for eligibility (Fig. 1). After exclusion criteria were applied, 17 studies provided data for the meta-analysis (i.e., 9 studies of sleep duration, 4 of sleep quality and 4 of both measures: providing 23 and 10 units of analysis for sleep duration and sleep quality, respectively) (Table 1; Fig. 1). Here we define units of analysis as the independent samples that enabled us to obtain any heritability estimate. As several studies reported heritability data for more than one sample, the units of analysis were slightly larger than the number of studies included in the meta-analyses.

### 2.2. Data extraction

For all the studies the following characteristics of the studies were extracted:

- Average age of the study sample with standard deviation
- Proportion of males and females
- Proportion of monozygotic (MZ), dizygotic (DZ), dizygotic opposite-sex (DZOS) and non-twin siblings (if relevant)
- Country of origin of the study population
- Continent of origin of the study population
- Type of measure: assessment using an objective/subjective measure,

**Table 1**
Main characteristics of studies included in the meta-analyses.

| | Study | Quality (Units of analysis) | Duration (Units of analysis) | Measure | Continent (Country) | Age | %Males | $H^2$ Sleep Quality | $H^2$ Sleep Duration |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Gedda and Brenci, 1979 | NO | YES (2) | Single question | Europe (Italy) | A: 6-8<br>B: 16-18 | A:48<br>B:46 | A: /<br>B: / | A:0<br>B:.23 |
| 2 | Partinen et al., 1983 | NO | YES (8) | Single question | Europe (Finland) | A: 18-24<br>B: 18-24<br>C: > 25<br>D: > 25<br>E: 18-24<br>F: 18-24<br>G: > 25<br>H: > 25 | A:100<br>B:0<br>C:100<br>D:0<br>E:100<br>F:0<br>G:100<br>H:0 | A: /<br>B: /<br>C: /<br>D: /<br>E: /<br>F: /<br>G: /<br>H: / | A: .20<br>B: .28<br>C: 1<br>D: .72<br>E: 0<br>F: .36<br>G: .32<br>H: .38 |
| 3 | Heath et al., 1990 | YES (1) | YES (1) | Sleep questionnaire (Johns and Paler) | Oceania (Australia) | 17-88 | 36 | .32 | .40 |
| 4 | de Castro, 2002 | NO | YES (1) | 7-day diary | America (USA) | $\bar{X} = 42$ | 52 | / | .30 |
| 5 | Boomsma et al., 2008 | YES (2) | YES (2) | Sleep questionnaire (Dutch Groningen Sleep Questionnaire) | Europe (The Netherlands) | A: $\bar{X} = 31$<br>B: $\bar{X} = 31$ | A = 100<br>B = 0 | A: 0<br>B: 0 | A: 0<br>B: 0 |
| 6 | Paunio et al., 2009 | YES (2) | NO | Single question | Europe (Finland) | A: $\bar{X} = 39$<br>B: $\bar{X} = 39$ | A = 100<br>B = 0 | A: .39<br>B: .39 | A: /<br>B: / |
| 7 | Barclay et al., 2010b | NO | YES (1) | Sleep questionnaire (PSQI) | Europe (United Kingdom) | $\bar{X} = 20$ | 38 | / | 0 |
| 8 | Barclay et al., 2010a | YES (1) | NO | Sleep questionnaire (PSQI) | Europe (United Kingdom) | $\bar{X} = 20$ | 38 | .43 | / |
| 9 | Watson et al., 2010 | NO | YES (1) | Single question | America (USA) | $\bar{X} = 37$ | 31 | / | .31 |
| 10 | Liu et al., 2012 | NO | YES (2) | Sleep questionnaire (PSQI) | Asia (China) | A: $\bar{X} = 31$<br>B: $\bar{X} = 39$ | A = 100<br>B = 0 | A: /<br>B: / | A: .27<br>B: .29 |
| 11 | Genderson et al., 2013 | YES (1) | YES (1) | Sleep questionnaire (PSQI) | America (USA) | $\bar{X} = 55$ | 100 | .34 | .29 |
| 12 | Sletten et al., 2013 | NO | YES (1) | Actigraphy and diary | Oceania (Australia) | $\bar{X} = 12$ | 37 | / | .65 |
| 13 | Butkovic et al., 2014 | NO | YES (1) | Single question | Europe (Croatia) | $\bar{X} = 19$ | / | / | .63 |
| 14 | Taylor et al., 2015 | YES (1) | NO | Sleep questionnaire (PSQI) | Europe (United Kingdom) | $\bar{X} = 16$ | 45 | .41 | / |
| 15 | Gasperi et al., 2017 | YES (1) | NO | Sleep questionnaire (PSQI) | America (USA) | $\bar{X} = 29$ | 37 | .36 | / |
| 16 | Madrid-Valero et al., 2018 | YES (1) | YES (1) | Sleep questionnaire (PSQI) | Europe (Spain) | $\bar{X} = 54$ | 45 | .34 | .30 |
| 17 | Rusterholz et al., 2018 | NO | YES (1) | Sleep EEG | Europe (Switzerland) | $\bar{X} = 12$ | 50 | / | 0.02 |
| | Total | 10 | 23 | | | | | | |

Note: Gedda et al., provided two units of analysis one for participants from 6 to 8 years old and the other for participants between 16 and 18 years old; Partinen et al., provided eight units of analyses based on: sex (male, female); age (18-24, ≥25) and cohabitation (yes,no), Boomsma et al., provided two units of analysis one for male and the other for females

single question/validated instrument. Whether sleep duration was categorized or assessed as a continuous variable
- Best fitting model (ACE, ADE, AE, CE or E)
- Number of twin/siblings pairs in the study
- Number of participants in the study
- MZ and DZ correlations
- Components of the variance for the full model: we entered the heritability ($h^2$) and shared environmental component ($c^2$) under the ACE or ADE model. When the reported model was an ACE model, the estimate for A was entered in $h^2$_FULL and C was entered in $c^2$_FULL. When an ADE model was reported, we summed A and D and entered the sum in $h^2$_FULL and zero for $c^2$_FULL. If both univariate and multivariate analyses were presented, estimates from univariate models were included
- Components of the variance for the BEST model: When the best fitting model was an ACE model, we entered A for $h^2$_BEST and C was entered for $c^2$_FULL. When the best fitting model was an ADE, we entered the sum of A and D in $h^2$_BEST and 0 in $c^2$_BEST. If the best fitting model was an AE (or CE or E) we entered zero for the component (or components) dropped and the significant components were entered as previously described

For computational reasons it is routine practice that correlations and heritability values of 1 or -1 be truncated to 0.999 or -0.999, respectively. When, due to the application of Falconer's formulae to atypical correlation patterns, the heritability value was reported as higher than 1 or lower than 0 it was truncated to 0.999 or 0, respectively.

Data entry checks: To assess the reliability of the data extraction process of the study characteristics, all the studies were double coded by J.J.M.V and M.R.A and disagreements were resolved by consensus with a third reviewer (J.S.M or J.O.M.). The results showed very satisfactory interrater reliability, with kappa coefficients ranging from .95 to 1 ($M$ = .97) for the categorical variables and intraclass correlations between 0.99 and 1 ($M$ = .99) for the continuous variables.

### 2.3. Statistical analyses

In this meta-analysis, the effect sizes were monozygotic and dizygotic twin correlations (rMZ and rDZ, respectively), and estimates of heritability from the full and best fitting models ($h^2$_FULL and $h^2$_BEST, respectively). The effect sizes were transformed into the Fisher's Z metric in order to normalize distributions and stabilize variances.

Separate meta-analyses were carried out for rMZ, rDZ, $h^2$_FULL and $h^2$_BEST for both key outcomes (i.e. sleep quality and sleep duration). In the case of sleep duration, some studies provided data from the full model but not for the model of best fit. In such cases heritability estimated from the full model were included in the best fitting model meta-analysis. Heritability from the full model would be in any case more conservative than that of the best model. Therefore, this procedure

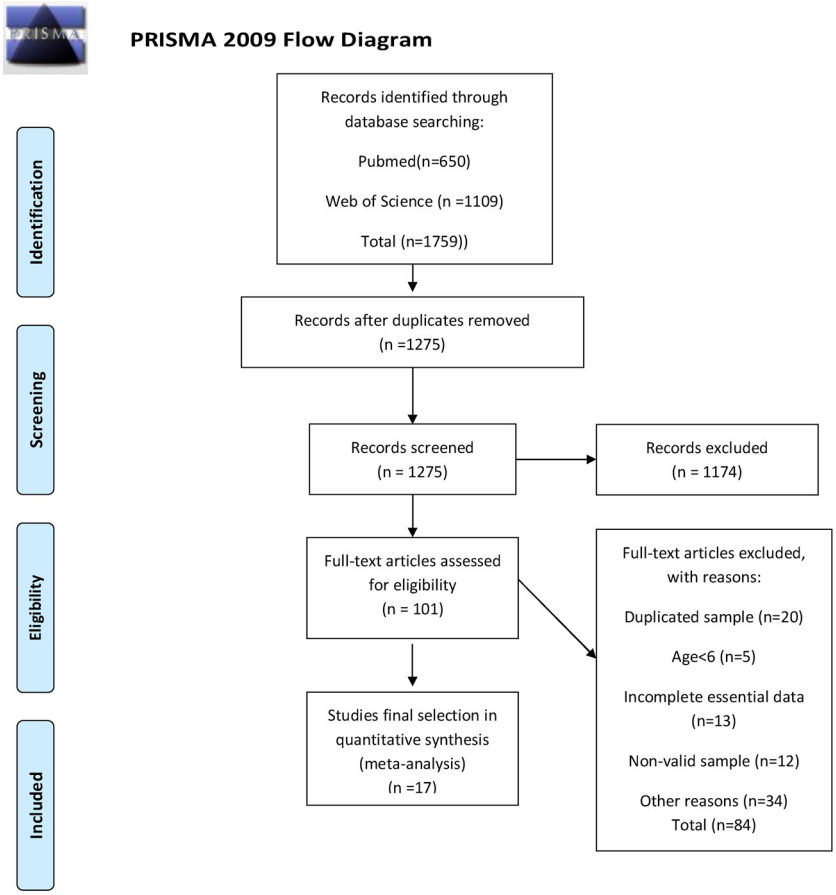


**Fig. 1.** Flow Chart of study selection process.

should not overestimate genetic influences. Thus, a total of 9 meta-analyses were conducted.

In order to calculate summary statistics of effect sizes in each of the aforementioned meta-analyses, a random-effects model was assumed (Borenstein et al., 2009; Sánchez-Meca et al., 2013). This model implies that each effect size is weighted by its inverse variance, which is the sum of the within-study variance and the between-studies variance. The between-studies variance was estimated by restricted maximum likelihood (López-López et al., 2013). In addition, 95% confidence intervals around each mean effect size were computed using the improved method proposed by Hartung (1999; Sanchez-Meca & Marin-Martinez, 2008). To facilitate the interpretation of the results, the mean effect sizes and their confidence limits (calculated on Fisher's Z transformed effect sizes) were back transformed into Pearson correlation metric.

For each meta-analysis, a forest plot, $Q$ statistic and $I^2$ index were used to assess heterogeneity among the effect sizes. The $Q$ statistic was applied to test the homogeneity assumption among the effect sizes by assuming a chi-square distribution with degrees of freedom equal to k – 1, k being the number of studies. A $p$ value for the Q statistic of $p < .05$ allows us to reject the homogeneity hypothesis (Borenstein et al., 2009). The $I^2$ index quantifies the degree of true heterogeneity exhibited by the effect sizes. Values of about 25%, 50%, and 75% reflect low, moderate, and large heterogeneity, respectively (Higgins et al., 2003).

For meta-analyses with at least 10 effect sizes, where evidence of heterogeneity was found, moderator analyses were performed assuming mixed-effects models. Weighted ANOVAs and meta-regressions were applied for categorical and continuous moderators, respectively. To test the statistical significance of each moderator variable, an $F$ test based on the improved method proposed by Knapp and Hartung was applied (Knapp & Hartung, 2003; López-López et al., 2013; Viechtbauer et al., 2015). $Q_E$ and $Q_w$ statistics were also applied for testing the model misspecification for meta-regressions and ANOVAs, respectively. The proportion of variance accounted for by the moderator variable was estimated with $R^2$ (López-López et al., 2014). Finally, the risk of publication bias was assessed with the Egger test and by constructing funnel plots with the trim-and-fill method to impute missing effect sizes in case of asymmetry in the funnel plot (Duval & Tweedie, 2000). All statistical analyses were carried out with *metafor* package in *R* (Viechtbauer, 2010).

## 3. Results

### 3.1. Description of the study designs and samples

Most of the studies applied the classical twin design which involves comparing MZ and DZ twin resemblance using correlation or variance/covariance matrices. One study included only men (Genderson et al., 2013) since participants belonged to the Vietnam Era Twin Study of Aging. Three studies from two different samples also included non-twin siblings in the analyses (Barclay et al., 2010a,b; Boomsma et al., 2008). One study (Paunio et al., 2009) reported on the heritability of sleep quality at two different time points (1975 and 1981). In this study the most recent value was selected.

### 3.2. Subjective sleep quality

Data were available in 5 units of analysis for rMZ and rDZ correlations and heritability estimated from the full model, and for all units of analysis (10 units of analysis) for heritability estimated from the best fitting model. The forest plot for heritability from the best fitting models is displayed in Fig. 2. The mean effect size for rMZ was $r_+ = .37$

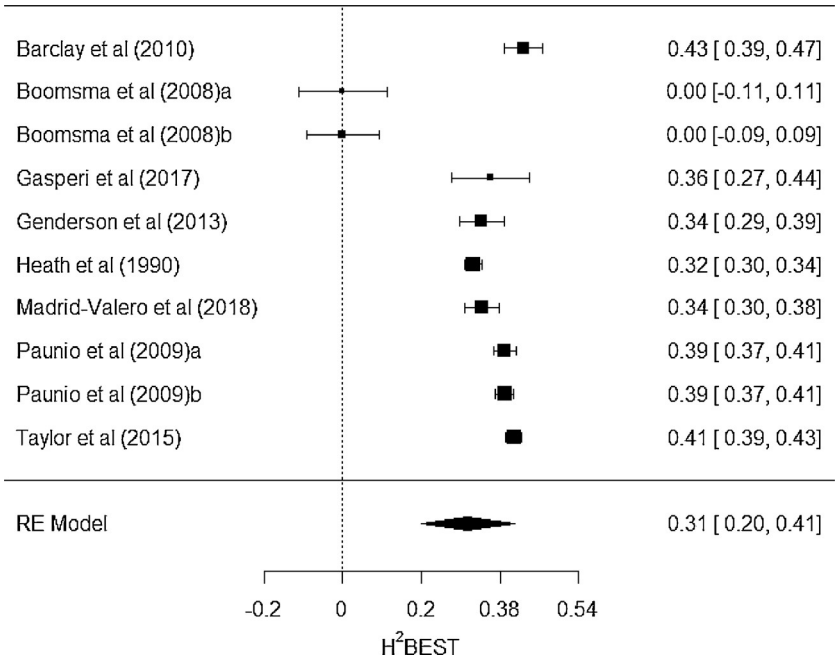


**Fig. 2.** Forest plot displaying the heritability estimates from the best fitting models (with 95% confidence intervals) for subjective sleep quality. Note. H2 BEST refers to the estimate provided from the model of best fit reported in the paper. RE model refers to the statistical model assumed in the calculations: Random-Effects Model.

(95% CI = .30, .44) and $r_+$ = .18 (95% CI = .13, .23) for rDZ. Regarding the MZ twin correlation, significant heterogeneity was found ($Q$(4) = 45.39, $p$ < 0.0001; $I^2$ = 86.49). However, there was a non-significant trend for heterogeneity for the DZ twin correlation ($Q$(4) = 8.39, $p$ = 0.078; $I^2$ = 55.23). Forest plots for MZ and DZ correlations are displayed in Supplementary Figures 1 & 2.

Similar mean effect sizes were found for heritability in the full ($r_+$ = .36; 95% CI = .31, .42) and best fitting ($r_+$ = .31; 95% CI = .20, .41) models. Both heterogeneity indices were significant (Table 2). Heritability values from the best fitting models ranged from 0 (Boomsma et al., 2008) to .43 (Barclay et al., 2010a). In the full model heritability estimates ranged from .32 (Heath et al., 1990) to .43 (Barclay et al., 2010a). The shared environment estimate was zero in both the best and full models.

A sensitivity analysis for the best fitting model was performed. Here, two units of analysis from Boomsma et al. (2008) were removed, since no genetic variance was found for males or females. The results showed a similar mean effect size (.37) and the heterogeneity index remained significant (95% CI = .34, .41; $Q$(7) = 61.86, $p$ < 0.0001; $I^2$ = 87.19).

**Table 2**
Average effect sizes, 95% confidence intervals, and heterogeneity statistics of twin correlations and variance components for subjective sleep quality outcome.

| | | | 95% CI | | | | |
|---|---|---|---|---|---|---|---|
| Statistic | $k$ | $r_+$ | LL | UL | $Q$ | $p$ | $I^2$ |
| $r_{MZ}$ | 5 | .372 | .302 | .438 | 45.392 | < .0001 | 86.49 |
| $r_{DZ}$ | 5 | .177 | .125 | .225 | 8.386 | .0784 | 55.23 |
| $h^2$_FULL | 5 | .363 | .307 | .417 | 23.511 | .0001 | 80.53 |
| $h^2$_BEST | 10 | .308 | .197 | .412 | 180.351 | < .0001 | 98.77 |

$r$ = correlation. MZ = monozygotic twins. DZ = dizygotic twins. $h^2$ _FULL = heritability estimates from the full model. $h^2$ _BEST = heritability estimates from the best fitting model. $k$ = number of units of analysis. $r_+$ = average effect size estimate. LL and UL: lower and upper limits of the 95% confidence interval (95% CI) for $r_+$. $Q$ = Cochran's heterogeneity $Q$ statistic; $Q$ statistic has $k$ – 1 degrees of freedom. $p$ = probability level for the $Q$ statistic. $I^2$ = heterogeneity index.

### 3.3. Analysis of publication bias (Subjective sleep quality)

Non-significant results for the interception were obtained from the Egger test for MZ and DZ twin correlations [$t$(3) = -0.25, $p$ = .818 and $t$ (3) = 0.23, $p$ = .834; respectively) as well as heritability estimates from the full model [$t$(3) = 0.46, $p$ = .678]. Supplementary Fig. 3 presents the funnel plots obtained with the original 5 MZ and 5 DZ twin correlations, as well as with the 5 heritability estimates from the full model (A, B, and C). Applying the trim-and-fill method (which involves imputing missing effect sizes to achieve symmetry in the funnel plot where necessary), no effect sizes had to be imputed. These results led us to discard publication bias as a threat against these meta-analytic results.

For the heritability estimates from the best fitting model, a significant result for the interception was obtained with the Egger test: $t$(8) = -2.97, $p$ = .018. By applying the trim-and-fill method, two additional heritability estimates were imputed to the set of original estimates to achieve symmetry in the funnel plot (see supplementary Fig. 3-D). When a mean effect (and its 95% CI) was calculated using the 10 heritability estimates plus the two imputed values, the average effect was $r_+$ = .28 (95% CI = .19, .36). If we compare the new effect with that obtained using the 10 original heritability estimates ($r_+$ = .31; 95% CI = .19, .41) only slight differences were found.

### 3.4. Sleep duration

For sleep duration data were available for 20 units of analysis, both for rMZ and rDZ correlations. Regarding heritability, 19 units of analysis provided heritability for the full model and 11 for the best fitting model. The mean effect size for rMZ was $r_+$ = .51 (95% CI = .34, .64) and $r_+$ = .29 (95% CI = .16, .40) for rDZ. There was significant heterogeneity for both correlations (rMZ: $Q$(19) = 501.98, $p$ < 0.0001; $I^2$ = 99.18; rDZ: $Q$(19) = 244.65, $p$ < 0.0001; $I^2$ = 97.84) (Table 3). Forest plots for MZ and DZ correlations are displayed in Supplementary Figures 4 & 5.

The heritability of sleep duration from the study units included in the meta-analysis ranged from 0 (Barclay et al., 2010b; Boomsma et al., 2008; Gedda & Brenci, 1979; Partinen et al., 1983) to 1 (Partinen et al., 1983). Regarding the mean effect sizes for heritability a value of $r_+$ =

**Table 3**
Average effect sizes, 95% confidence intervals, and heterogeneity statistics of twin correlations and variance components for sleep duration outcome.

| | | | 95% CI | | | | |
|---|---|---|---|---|---|---|---|
| Statistic | *k* | $r_+$ | LL | UL | *Q* | *p* | $I^2$ |
| $r_{MZ}$ | 20 | .508 | .343 | .644 | 501.98 | < .0001 | 99.18 |
| $r_{DZ}$ | 20 | .286 | .162 | .402 | 244.65 | < .0001 | 97.84 |
| $h^2$_FULL | 19 | .410 | .152 | .616 | 2232.83 | < .0001 | 99.79 |
| $h^2$_BEST | 11 | .301 | .133 | .452 | 463.33 | < .0001 | 98.78 |
| $h^2$_MIXED | 23 | .379 | .159 | .562 | 2301.02 | < .0001 | 99.73 |

*r* = correlation. MZ = monozygotic twins. DZ = dizygotic twins. $h^2$ _FULL = heritability estimates from the full model. $h^2$ _BEST = heritability estimates from the best fitting model. $h^2$ _MIXED = heritability estimates from the full and best fitting model. *k* = number of units of analysis. $r_+$ = average effect size estimate. LL and UL: lower and upper limits of the 95% confidence interval (95% CI) for $r_+$. *Q* = Cochran's heterogeneity *Q* statistic; *Q* statistic has *k* – 1 degrees of freedom. *p* = probability level for the *Q* statistic. $I^2$ = heterogeneity index.

.41 (95% CI = .15, .62) was found from the full models and $r_+$ = .30 (95% CI = .13, .45) from the best fitting models. As some studies did not provide data for the full model or the best fitting model, we combined estimates from the full models with those from the best fitting models (i.e. heritability estimates from the full model were also assumed as the model of best fit, when only the full model was reported) in order to meta-analyse all the studies at the same time. Results from this mixed model showed a mean effect size of $r_+$ = .38 (95% CI = .16, .56) (Fig. 3). Only two studies reported significant shared environmental influences, .23 (CIs were not provided) (Boomsma et al., 2008) and .26 (95% CI = 0, .37) (b).

Given the heterogeneity of the measures used to assess sleep duration and the possibility that there could be different heritability estimates for self-report and objective measures, we performed a sensitivity analysis focusing on self-report data exclusively. The results showed no substantial change for the mean effect size or the heterogeneity index [Best fitting model: r+ = .30 (95% CI = .13, .45); Q(10) = 463.33, $p < 0.0001$; $I^2$ = 98.78; Mixed model: r+ = .39 (95% CI = .17, .58); Q (21) = 2294.46, $p < 0.0001$; $I^2$ = 99.75].

### 3.5. Analysis of publication bias (Sleep duration)

Significant results for the interception were obtained using the Egger test for MZ [$t(18) = 4.73$, $p < .001$] and DZ [$t(18) = 3.77$, $p = .001$] twin correlations. Supplementary Figure 6 (A and B) presents the funnel plots obtained with the original 20 MZ and DZ twin correlations. Applying the trim-and-fill method, no effect sizes had to be imputed to achieve symmetry in any of the funnel plots.

Regarding heritability estimates, non-significant results for the interception were obtained using the Egger test in every case [Full model: $t(17) = 0.17$, $p = .866$; best fitting model: $t(9) = 0.83$, $p = .429$; full + best fitting model: $t(21) = 0.21$, $p = .838$]. Supplementary Figure 6 (C, D, and E) presents the funnel plots obtained with the original heritability estimates from the full, best and full + best models. Applying the trim-and-fill method, no effect sizes had to be imputed to achieve symmetry in any of the funnel plots. These results led us to discard publication bias as a threat against these meta-analytic results.

### 3.6. Analyses of moderator variables

A number of moderators of the results were considered: mean age of the participants in the sample, SD of the sample age, % MZ twin pairs, % DZOS twin pairs, continent of origin, and type of measure. None of them showed a statistically significant relationship with the effect size in any of the models (best, full or best + full in the case of duration). The results of these moderator analyses are reported in Supplementary Tables 2, 3, 4 and 5 for the heritability of subjective sleep quality (best fitting model) and sleep duration (best + full model).

## 4. Discussion

This meta-analysis examined the genetic and environmental influences on two general measures of sleep. Our results show that genetic factors account for a significant proportion of the phenotypic variance in both measures. Mean effect sizes for heritability were similar and of moderate magnitude for subjective sleep quality and sleep duration [.31 (.20 - .41) and .38 (.16 -.56) respectively]. The presence of a significant heritability is in line with most of the studies included in the meta-analysis since only a few did not find any genetic influence for

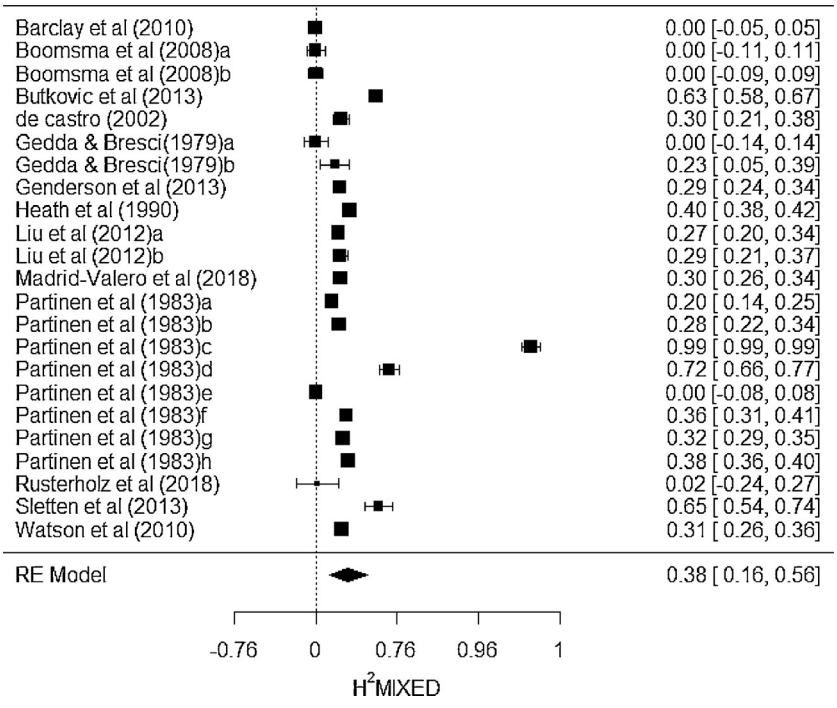


**Fig. 3.** Forest plot displaying the heritability estimates from the full and best fitting models (and 95% confidence intervals) for sleep duration. Note H2 MIXED refers to data from the model of best fit (except where this was not reported, in which case data from the full model are used). RE model refers to the statistical model assumed in the calculations: Random-Effects Model.

sleep quality (Boomsma et al., 2008) or sleep duration (Barclay et al., 2010b; Boomsma et al., 2008; Gedda & Brenci, 1979, Partinen et al., 1983). Additionally, all heritability analyses showed significant heterogeneity, pointing towards high variability across different studies. These discrepancies could be explained by sample characteristics, including sample size, age, or sex distribution. They could also be associated with population-specific environmental issues, such as cultural background or employment situation; or by methodological issues, such as the instruments/techniques used to assess sleep measures.

Our results also highlight the role of unique environmental factors (estimates which include measurement error) since they explain the largest proportion of the variance in most of the studies. Furthermore, none of the reports analyzed for sleep quality found a significant impact of shared environmental factors and only two studies for sleep duration did (and for those studies, shared environmental influences were low). These results chime well with twin studies focusing on other phenotypes, where C is usually small and decreases with age to the point of becoming undetectable (Plomin et al., 2016).

Supporting the substantial role of genetic factors in the phenotypic variance of sleep-related variables, several GWAS have explored gene-sets and pathways involved in different measures, including self-reported sleep duration, insomnia, daytime sleepiness, or chronotype, and are detecting an increasing number of loci implicated (Dashti et al., 2019; Doherty et al., 2018; Jansen et al., 2019; Jones et al., 2019; Jones et al., 2016; Kalmbach et al., 2017; Lane et al., 2017; Nishiyama et al., 2019). Some consistent replications for sleep duration or insomnia in PAX-8, VRK2 or MEIS-1 (Dashti et al., 2019; Doherty et al., 2018; Jones et al., 2016; Lane et al., 2017) have been reported, together with many other genes and genetic correlations with other health-relevant phenotypes, including anthropometric, cognitive, metabolic and psychiatric traits (Dashti et al., 2019; Jansen et al., 2019). Multiple biological pathways appear to be involved in sleep variability and relationships with other variables: gene set enrichments for cortical and subcortical tissues, including striatum, hypothalamus or subpallium, mechanosensory response, dopamine binding or synaptic neurotransmission and plasticity have been reported (Dashti et al., 2019; Jansen et al., 2019). Altogether, the increasing power of GWAS is steadily getting closer to matching the results of quantitative and molecular genetics analyses.

As noted in the introduction, age and sex are both relevant variables for sleep quality and disturbances (Madrid-Valero et al., 2017a; Zhang & Wing, 2006). However, our meta-analysis did not detect significant moderating influences of these two variables for the heritability of sleep quality or duration. Although it is possible that age and sex were not important moderators, failure to find an association could also be due to a limited number of studies included in the meta-analysis. Moreover, most of the studies do not analyse age or sex effects specifically, in spite of using large samples with a wide age range. Regarding sex it seems that genetic influences do not differ for males and females since Madrid-Valero et al. (2018) found no evidence for sex differences in genetic influences on sleep quality; and Boomsma et al. (2008) found the same model and estimates for sleep quality and sleep duration in both males and females.

Although there is not an excessively large number of studies addressing the heritability of sleep quality and sleep duration (especially the former), it is noteworthy that these studies are published in journals of different scope. These journals include: sleep journals (SLEEP, Journal of Sleep Research), clinical journals (Psychosomatic Medicine, Journal of Abnormal Psychology), Epidemiological journals (American Journal of Epidemiology) and behavioral genetics journals (Twin Research and Human Genetics). This underscores the importance of twin studies in different fields and also the significance of sleep for general health.

This systematic review and the meta-analyses reported here demonstrate that genetic factors account for a substantial proportion of the variance for sleep quality and sleep duration. The systematic review also highlighted that genetic factors play a role in the association between these sleep variables and a wide variety of other variables relevant for general health. Twin studies have demonstrated that there is a substantial genetic overlap between sleep quality/duration/ sleep disorders and depression (Gasperi et al., 2017; Gehrman et al., 2011; Gregory et al., 2011; Lind et al., 2017), anxiety (Gehrman et al., 2011; Gregory et al., 2011), psychotic-like experiences (Taylor et al., 2015), neuroticism (Butkovic et al., 2014), externalising behaviors (Barclay et al., 2011; Madrid-Valero et al., 2019) and pain (Gasperi et al., 2017; Pinheiro et al., 2018) among other variables relevant for health. These results are largely consistent across the behavioral genetic literature and point to a significant genetic correlation between sleep disturbances and different disorders or diseases. Altogether, this means that genetic factors explaining individual differences in the way that we sleep may be related to each one of those traits, and the possibility remains that some gene-set or pathway serving a basal function is common to most or all of them. In line with this, GWAS studies have also found large genetic correlations between insomnia and anxiety, depressive symptoms, subjective well-being, and neuroticism among others (Jansen et al., 2019).

Twin studies are less subject to publication bias as compared to other scientific reports since finding a low heritability is as interesting as finding a high heritability. Twin samples are usually large, and effect sizes (heritability) typically explain between 30-50% of the variance (Plomin et al., 2016). Despite the robustness and methodological strengths of twin studies there are a paucity of studies that address the heritability of sleep quality or sleep duration using comparable samples in terms of age, sex and cultural/geographical origins. In addition, there is a need for consensus in the definition and measurement of the main sleep phenotypes, as well as further well-powered studies including objective measures of sleep. As heritability may vary by sample and phenotypic characteristics, research should go beyond comparing twins in studies to report upon heritability and instead search for comparability and generalizability of their results. Authors should be encouraged to replicate investigations in different populations and analyse possible variations due to sample, measurement or environmental characteristics. There is a significant heterogeneity across studies and further research is needed to elucidate reasons for this variability.

### 4.1. Strengths and limitations

This study has several strengths. The focus on two different general aspects of sleep allows us to have greater insight into genetic and environmental influences on individual differences for these phenotypes. Additionally, twin samples are usually large and, as stated before, these kind of studies typically show large effect sizes and are less subject to publication bias as compared to other scientific studies (Plomin et al., 2016). However, results from this systematic review and meta-analysis should also be interpreted in light of some limitations. For example, the small number of studies included in this report and the non-comparability of the samples did not allow us to find moderator variables which could explain the heterogeneity reported among studies. Additionally, most of the studies included in this report used subjective measures to assess sleep duration - and all of them used subjective measures to assess sleep quality. While the authors considered this focus to be the optimal approach for the current meta-analysis, further studies using well-powered objective measures are needed to investigate possible differences between objective and subjective measures of sleep. Moreover, some studies did not report twin correlations or heritability estimates from both full and best fitting models. Reports from both models and correlations should be included in all papers reporting classic twin studies. Finally, the methods used to estimate phenotypic variation and heritability could influence the results. In this meta-analysis different methods were used to calculate the parameters of the variance (e.g. Falconer's formulae or Structural Equation Modelling) and different instruments (e.g. the Pittsburgh Sleep Quality

Index or a single question) were used to assess sleep quality and duration.

### 4.2. Future directions

Future studies should address the heritability of sleep quality and sleep duration in different populations, using different measures, and also test if genetic influences vary across the life span. Further GWAS should continue addressing the specific gene-sets and pathways associated with sleep quality and sleep duration. This knowledge might be of use when developing tailored treatments. Polygenic Risk Scores could be used to identify people at risk of developing sleep difficulties – providing an opportunity to limit the progression of these problems and their impact on associated traits such as depression or anxiety. Finally, this meta-analysis has highlighted the need for further research regarding the genetic and environmental influences of sleep quality and sleep duration since there is substantial heterogeneity across studies and we have not been able to identify moderators underlying these differences.

### 4.3. Conclusions

Approximately one third of the variation in subjective sleep quality and sleep duration scores, in general population samples, is explained by genetic factors. Moreover, there is a substantial heterogeneity across studies. Therefore, further research is needed to identify moderators underlying these differences.

## Declaration of Competing Interest

Alice Gregory is an advisor for a project sponsored by Johnson's Baby. She has written two books (Nodding Off, Bloomsbury Sigma, 2018; The Sleepy Pebble, Flying Eye Books, 2019). She is a regular contributor to BBC Focus magazine and has contributed to numerous other outlets (such as The Conversation and The Guardian). She has been interviewed by magazines and commercial websites. She has provided a talk for business (Investec) and is occasionally sent trial products from commercial companies (e.g. blue light blocking glasses).

## Acknowledgments

J.J.M-V. was supported by predoctoral scholarship (19814/FPI/15) of the Fundación Seneca.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.neubiorev.2019.12.028.

## References[1]


Aili, K., Åström-Paulsson, S., Stoetzer, U., Svartengren, M., Hillert, L., 2017. Reliability of Actigraphy and Subjective Sleep Measurements in Adults: The Design of Sleep Assessments. J Clin Sleep Med 13 (1), 39–47. https://doi.org/10.5664/jcsm.6384.

Alvaro, P.K., Roberts, R.M., Harris, J.K., 2013. A Systematic Review Assessing Bidirectionality between Sleep Disturbances, Anxiety, and Depression. Sleep 36 (7), 1059–1068. https://doi.org/10.5665/sleep.2810.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. American Psychiatric Association, Arlington, VA.

Barclay*, N.L., Eley, T.C., Buysse, D.J., Archer, S.N., Gregory, A.M., 2010a. Diurnal preference and sleep quality: same genes? A study of young adult twins. Chronobiol Int 27 (2), 278–296. https://doi.org/10.3109/07420521003663801.

Barclay*, N.L., Eley, T.C., Buysse, D.J., Rijsdijk, F.V., Gregory, A.M., 2010b. Genetic and environmental influences on different components of the Pittsburgh Sleep Quality Index and their overlap. Sleep 33 (5), 659–668.

Barclay, N.L., Eley, T.C., Maughan, B., Rowe, R., Gregory, A.M., 2011. Associations between diurnal preference, sleep quality and externalizing behaviors: a behavioral genetic analysis. Psychol Med 41 (5), 1029–1040. https://doi.org/10.1017/S0033291710001741.

Barclay, N.L., Gregory, A.M., 2013. Quantitative genetic research on sleep: a review of normal sleep, sleep disturbances and associated emotional, behavioral, and health-related difficulties. Sleep Med Rev 17 (1), 29–40. https://doi.org/10.1016/j.smrv.2012.01.008.

Bokenberger, K., Ström, P., Dahl Aslan, A.K., Johansson, A.L., Gatz, M., Pedersen, N.L., Åkerstedt, T., 2017. Association Between Sleep Characteristics and Incident Dementia Accounting for Baseline Cognitive Status: A Prospective Population-Based Study. J Gerontol A Biol Sci Med Sci 72 (1), 134–139. https://doi.org/10.1093/gerona/glw127.

Boomsma*, D.I., van Someren, E.J., Beem, A.L., de Geus, E.J., Willemsen, G., 2008. Sleep during a regular week night: a twin-sibling study. Twin Res Hum Genet 11 (5), 538–545. https://doi.org/10.1375/twin.11.5.538.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. Introduction to meta-analysis. Wiley, Chichester, UK.

Butkovic*, A., Vukasovic, T., Bratko, D., 2014. Sleep duration and personality in Croatian twins. J Sleep Res 23 (2), 153–158. https://doi.org/10.1111/jsr.12101.

Buysse, D.J., Hall, M.L., Strollo, P.J., Kamarck, T.W., Owens, J., Lee, L., et al., 2008. Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical/polysomnographic measures in a community sample. J Clin Sleep Med 4 (6), 563–571.

Cappuccio, F.P., D'Elia, L., Strazzullo, P., Miller, M.A., 2010. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep 33 (5), 585–592. https://doi.org/10.1093/sleep/33.5.585.

Cirelli, C., Tononi, G., 2008. Is sleep essential? Plos Biology 6 (8), 1605–1611. https://doi.org/10.1371/journal.pbio.0060216.

Dashti, H.S., Jones, S.E., Wood, A.R., Lane, J.M., van Hees, V.T., Wang, H., et al., 2019. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. Nat Commun 10 (1), 1100. https://doi.org/10.1038/s41467-019-08917-4.

de Castro*, J.M., 2002. The influence of heredity on self-reported sleep patterns in free-living humans. Physiol Behav 76 (4-5), 479–486.

Desai, A.V., Cherkas, L.F., Spector, T.D., Williams, A.J., 2004. Genetic influences in self-reported symptoms of obstructive sleep apnoea and restless legs: a twin study. Twin Res 7 (6), 589–595. https://doi.org/10.1375/1369052042663841.

Doherty, A., Smith-Byrne, K., Ferreira, T., Holmes, M.V., Holmes, C., Pulit, S.L., Lindgren, C.M., 2018. GWAS identifies 14 loci for device-measured physical activity and sleep duration. Nat Commun 9 (1), 5257. https://doi.org/10.1038/s41467-018-07743-4.

Gasperi*, M., Herbert, M., Schur, E., Buchwald, D., Afari, N., 2017. Genetic and Environmental Influences on Sleep, Pain, and Depression Symptoms in a Community Sample of Twins. Psychosom Med 79 (6), 646–654. https://doi.org/10.1097/PSY.0000000000000456.

Gedda*, L., Brenci, G., 1979. Sleep and dream characteristics in twins. Acta Genet Med Gemellol (Roma) 28 (3), 237–239.

Gehrman, P.R., Ghorai, A., Goodman, M., McCluskey, R., Barilla, H., Almasy, L., et al., 2019. Twin-based heritability of actimetry traits. Genes Brain Behav 18 (5), e12569. https://doi.org/10.1111/gbb.12569.

Gehrman, P.R., Meltzer, L.J., Moore, M., Pack, A.I., Perlis, M.L., Eaves, L.J., Silberg, J.L., 2011. Heritability of insomnia symptoms in youth and their relationship to depression and anxiety. Sleep 34 (12), 1641–1646. https://doi.org/10.5665/sleep.1424.

Genderson*, M.R., Rana, B.K., Panizzon, M.S., Grant, M.D., Toomey, R., Jacobson, K.C., et al., 2013. Genetic and environmental influences on sleep quality in middle-aged men: a twin study. Journal of Sleep Research 22 (5), 519–526. https://doi.org/10.1111/jsr.12048.

Grandner, M.A., 2019. Epidemiology of insufficient sleep and poor sleep quality. In: Grandner, M.A. (Ed.), Sleep and Health. Academic Press, London, pp. 11–20.

Gregory, A.M., Agnew-Blais, J.C., Matthews, T., Moffitt, T.E., Arseneault, L., 2017. ADHD and Sleep Quality: Longitudinal Analyses From Childhood to Early Adulthood in a Twin Cohort. Journal of Clinical Child and Adolescent Psychology 46 (2), 284–294. https://doi.org/10.1080/15374416.2016.1183499.

Gregory, A.M., Buysse, D.J., Willis, T.A., Rijsdijk, F.V., Maughan, B., Rowe, R., et al., 2011. Associations between sleep quality and anxiety and depression symptoms in a sample of young adult twins and siblings. In *J Psychosom Res* 71, 250–255 England: 2011 Elsevier Inc.

Gregory, A.M., Rijsdijk, F.V., Eley, T.C., Buysse, D.J., Schneider, M.N., Parsons, M., Barclay, N.L., 2016. A Longitudinal Twin and Sibling Study of Associations between Insomnia and Depression Symptoms in Young Adults. Sleep 39 (11), 1985–1992. https://doi.org/10.5665/sleep.6228.

Gregory, A.M., Sadeh, A., 2012. Sleep, emotional and behavioral difficulties in children and adolescents. Sleep Med Rev Vol. 16. A 2011 Elsevier Ltd, England, pp. 129–136.

Gregory, A.M., Sadeh, A., 2016. Annual Research Review: Sleep problems in childhood psychiatric disorders–a review of the latest science. J Child Psychol Psychiatry 57 (3), 296–317. https://doi.org/10.1111/jcpp.12469.

Hartung, J., 1999. An alternative method for meta-analysis. Biometrical Journal 41 (8), 901–916. https://doi.org/10.1002/(SICI)1521-4036(199912)41:8<901::AID-BIMJ901>3.3.CO;2-N.

Heath*, A.C., Kendler, K.S., Eaves, L.J., Martin, N.G., 1990. Evidence for genetic influences on sleep disturbance and sleep pattern in twins. Sleep 13 (4), 318–335. https://doi.org/10.1093/sleep/13.4.318.

Higgins, J., Thompson, S., Deeks, J., Altman, D., 2003. Measuring inconsistency in meta-analyses. British Medical Journal 327 (7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.

Hirshkowitz, M., Whiton, K., Albert, S.M., Alessi, C., Bruni, O., DonCarlos, L., et al., 2015.


[1] References with an asterisk were included in the meta-analyses

National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health 1 (1), 40–43. https://doi.org/10.1016/j.sleh.2014.12.010.
Hublin, C., Partinen, M., Koskenvuo, M., Kaprio, J., 2007. Sleep and mortality: A population-based 22-year follow-up study. Sleep 30 (10), 1245–1253. https://doi.org/10.1093/sleep/30.10.1245.
Hublin, C., Partinen, M., Koskenvuo, M., Kaprio, J., 2011. Heritability and mortality risk of insomnia-related symptoms: a genetic epidemiologic study in a population-based twin cohort. Sleep 34 (7), 957–964. https://doi.org/10.5665/SLEEP.1136.
Jackowska, M., Dockray, S., Hendrickx, H., Steptoe, A., 2011. Psychosocial factors and sleep efficiency: discrepancies between subjective and objective evaluations of sleep. Psychosom Med 73 (9), 810–816. https://doi.org/10.1097/PSY.0b013e3182359e77.
Jansen, P.R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A.R., et al., 2019. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet* 51 (3), 394–403. https://doi.org/10.1038/s41588-018-0333-3.
Jones, S.E., Lane, J.M., Wood, A.R., van Hees, V.T., Tyrrell, J., Beaumont, R.N., et al., 2019. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. Nat Commun 10 (1), 343. https://doi.org/10.1038/s41467-018-08259-7.
Jones, S.E., Tyrrell, J., Wood, A.R., Beaumont, R.N., Ruth, K.S., Tuke, M.A., et al., 2016. Genome-Wide Association Analyses in 128,266 Individuals Identifies New Morningness and Sleep Duration Loci. PLoS Genet 12 (8), e1006125. https://doi.org/10.1371/journal.pgen.1006125.
Kalmbach, D.A., Schneider, L.D., Cheung, J., Bertrand, S.J., Kariharan, T., Pack, A.I., Gehrman, P.R., 2017. Genetic Basis of Chronotype in Humans: Insights From Three Landmark GWAS. Sleep 40 (2). https://doi.org/10.1093/sleep/zsw048.
Knapp, G., Hartung, J., 2003. Improved tests for a random effects meta-regression with a single covariate. Statistics in Medicine 22 (17), 2693–2710. https://doi.org/10.1002/sim.1482.
Knopik, V.S., Neiderhiser, J.M., DeFries, J.C., Plomin, R., 2017. Behavioral genetics, 7th ed. Worth Publishers, New York.
Krystal, A.D., Edinger, J.D., 2008. Measuring sleep quality. Sleep Med 9 (Suppl 1), S10–17. https://doi.org/10.1016/S1389-9457(08)70011-X.
Lane, J.M., Liang, J., Vlasac, I., Anderson, S.G., Bechtold, D.A., Bowden, J., et al., 2017. Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. Nat Genet 49 (2), 274–281. https://doi.org/10.1038/ng.3749.
Lauderdale, D.S., Knutson, K.L., Yan, L.L., Liu, K., Rathouz, P.J., 2008. Self-reported and measured sleep duration: how similar are they? Epidemiology 19 (6), 838–845. https://doi.org/10.1097/EDE.0b013e318187a7b0.
Lessov-Schlaggar, C.N., Bliwise, D.L., Krasnow, R.E., Swan, G.E., Reed, T., 2008. Genetic association of daytime sleepiness and depressive symptoms in elderly men. Sleep 31 (8), 1111–1117.
Lewis, K.J.S., Richards, A., Karlsson, R., Leonenko, G., Jones, S.E., Jones, H.J., et al., 2019. Comparison of Genetic Liability for Sleep Traits Among Individuals With Bipolar Disorder I or II and Control Participants. JAMA Psychiatry. https://doi.org/10.1001/jamapsychiatry.2019.4079.
Lind, M.J., Aggen, S.H., Kirkpatrick, R.M., Kendler, K.S., Amstadter, A.B., 2015. A Longitudinal Twin Study of Insomnia Symptoms in Adults. Sleep 38, 1423–1430. https://doi.org/10.5665/sleep.4982.
Lind, M.J., Hawn, S.E., Sheerin, C.M., Aggen, S.H., Kirkpatrick, R.M., Kendler, K.S., Amstadter, A.B., 2017. An examination of the etiologic overlap between the genetic and environmental influences on insomnia and common psychopathology. Depress Anxiety 34 (5), 453–462. https://doi.org/10.1002/da.22587.
Liu*, R., Liu, X., Arguelles, L.M., Patwari, P.P., Zee, P.C., Chervin, R.D., et al., 2012. A Population-Based Twin Study on Sleep Duration and Body Composition. Obesity 20 (1), 192–199. https://doi.org/10.1038/oby.2011.274.
Liu, T.Z., Xu, C., Rota, M., Cai, H., Zhang, C., Shi, M.J., et al., 2017. Sleep duration and risk of all-cause mortality: A flexible, non-linear, meta-regression of 40 prospective cohort studies. Sleep Med Rev 32, 28–36. https://doi.org/10.1016/j.smrv.2016.02.005.
López-Mínguez, J., Colodro-Conde, L., Bandín, C., Ordoñana, J.R., Garaulet, M., Madrid, J.A., 2016. Application of multiparametric procedures for assessing the heritability of circadian health. Chronobiol Int 33 (2), 234–244. https://doi.org/10.3109/07420528.2015.1130051.
López-Mínguez, J., Morosoli, J.J., Madrid, J.A., Garaulet, M., Ordoñana, J.R., 2017. Heritability of siesta and night-time sleep as continuously assessed by a circadian-related integrated measure. Sci Rep 7 (1), 12340. https://doi.org/10.1038/s41598-017-12460-x.
López-Mínguez, J., Ordoñana, J.R., Sánchez-Romera, J.F., Madrid, J.A., Garaulet, M., 2015. Circadian system heritability as assessed by wrist temperature: a twin study. Chronobiol Int 32 (1), 71–80. https://doi.org/10.3109/07420528.2014.955186.
Lovato, N., Lack, L., 2019. Insomnia and mortality: A meta-analysis. Sleep Med Rev 43, 71–83. https://doi.org/10.1016/j.smrv.2018.10.004.
Luyster, F., Strollo, P., Zee, P., Walsh, J., Sleep, B.D.A.A., 2012. Sleep: A Health Imperative. Sleep 35 (6), 727–734. https://doi.org/10.5665/sleep.1846.
López-López, J., Botella, J., Sánchez-Meca, J., Marín-Martínez, F., 2013. Alternatives for Mixed-Effects Meta-Regression Models in the Reliability Generalization Approach: A Simulation Study. Journal of Educational and Behavioral Statistics 38 (5), 443–469. https://doi.org/10.3102/1076998612466142.
López-López, J.A., Marín-Martínez, F., Sánchez-Meca, J., Van den Noortgate, W., Viechtbauer, W., 2014. Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. Br J Math Stat Psychol 67 (1), 30–48. https://doi.org/10.1111/bmsp.12002.
Madrid-Valero, J.J., Barclay, N.L., Rowe, R., Rotem, P., Buysse, D.J., Ordoñana, J.R., et al., 2020. Association between symptoms of sleep apnea and problem behaviors in young adult twins and siblings. Psychological Medicine In press.
Madrid-Valero, J.J., Martínez-Selva, J.M., Ribeiro do Couto, B., Sánchez-Romera, J.F., Ordoñana, J.R., 2017a. Age and gender effects on the prevalence of poor sleep quality in the adult population. Gaceta Sanitaria 31 (1), 18–22. https://doi.org/10.1016/j.gaceta.2016.05.013.
Madrid-Valero, J.J., Martínez-Selva, J.M., Ordoñana, J.R., 2017b. Sleep quality and body mass index: a co-twin study. J Sleep Res 26 (4), 461–467. https://doi.org/10.1111/jsr.12493.
Madrid-Valero, J.J., Ordoñana, J.R., Klump, K.L., Burt, S.A., 2019. Children Sleep and Antisocial Behavior: Differential Association of Sleep with Aggression and Rule-Breaking. J Abnorm Child Psychol 47 (5), 791–799. https://doi.org/10.1007/s10802-018-0480-0.
Madrid-Valero*, J.J., Sánchez-Romera, J.F., Gregory, A.M., Martínez-Selva, J.M., Ordoñana, J.R., 2018. Heritability of sleep quality in a middle-aged twin sample from Spain. Sleep 41 (9), 1–7. https://doi.org/10.1093/sleep/zsy110.
Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., et al., 2009. Finding the missing heritability of complex diseases. Nature 461 (7265), 747–753. https://doi.org/10.1038/nature08494.
Mayhew, A.J., Meyre, D., 2017. Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. Curr Genomics 18 (4), 332–340. https://doi.org/10.2174/1389202918666170307161450.
Medda, E.A.G., Delfino, D., Fagani, C., Ferri, M., Violani, C., Stazi, M.A., 2019. Adolescents self-reported sleep quality and emotional regulation: a discordant twin study. Ann Ist Super Sanità 55, 118–123. https://doi.org/10.4415/ANN_19_02_03.
Meers, J., Stout-Aguilar, J., Nowakowski, S., 2019. Sex differences in sleep health. In: Grandner, M.A. (Ed.), Sleep and Health. Academic Press, London, pp. 21–29.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D., Grp, P., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Plos Medicine 6 (7). https://doi.org/10.1371/journal.pmed.1000097.
Mong, J.A., Cusmano, D.M., 2016. Sex differences in sleep: impact of biological sex and sex steroids. Philos Trans R Soc Lond B Biol Sci 371 (1688), 20150110. https://doi.org/10.1098/rstb.2015.0110.
Nishiyama, T., Nakatochi, M., Goto, A., Iwasaki, M., Hachiya, T., Sutoh, Y., et al., 2019. Genome-wide association meta-analysis and Mendelian randomization analysis confirm ALDH2 influencing on sleep duration in the Japanese population. Sleep 42 (6). https://doi.org/10.1093/sleep/zsz046.
Odintsova, V.V., Willemsen, G., Dolan, C.V., Hottenga, J.J., Martin, N.G., Slagboom, P.E., et al., 2018. Establishing a Twin Register: An Invaluable Resource for (Behavior) Genetic, Epidemiological, Biomarker, and' Omics' Studies. Twin Res Hum Genet 21 (3), 239–252. https://doi.org/10.1017/thg.2018.23.
Ohayon, M., Wickwire, E.M., Hirshkowitz, M., Albert, S.M., Avidan, A., Daly, F.J., et al., 2017. National Sleep Foundation's sleep quality recommendations: first report. Sleep Health 3 (1), 6–19. https://doi.org/10.1016/j.sleh.2016.11.006.
Ohayon, M.M., Carskadon, M.A., Guilleminault, C., Vitiello, M.V., 2004. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep 27 (7), 1255–1273.
Partinen*, M., Kaprio, J., Koskenvuo, M., Putkonen, P., Langinvainio, H., 1983. Genetic and environmental determination of human sleep. Sleep 6 (3), 179–185. https://doi.org/10.1093/sleep/6.3.179.
Paunio*, T., Korhonen, T., Hublin, C., Partinen, M., Kivimäki, M., Koskenvuo, M., Kaprio, J., 2009. Longitudinal study on poor sleep and life dissatisfaction in a nationwide cohort of twins. Am J Epidemiol 169 (2), 206–213. https://doi.org/10.1093/aje/kwn305.
Pinheiro, M.B., Morosoli, J.J., Ferreira, M.L., Madrid-Valero, J.J., Refshauge, K., Ferreira, P.H., Ordonana, J.R., 2018. Genetic and Environmental Contributions to Sleep Quality and Low Back Pain: A Population-Based Twin Study. Psychosomatic Medicine 80 (3), 263–270. https://doi.org/10.1097/psy.0000000000000548.
Plante, D.T., 2019. The Importance of Sleep Phenotypes in Bipolar Disorders. JAMA Psychiatry. https://doi.org/10.1001/jamapsychiatry.2019.3747.
Plomin, R., DeFries, J.C., Knopik, V.S., Neiderhiser, J.M., 2016. Top 10 Replicated Findings From Behavioral Genetics. Perspect Psychol Sci 11 (1), 3–23. https://doi.org/10.1177/1745691615617439.
Polderman, T.J., Benyamin, B., de Leeuw, C.A., Sullivan, P.F., van Bochoven, A., Visscher, P.M., Posthuma, D., 2015. Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat Genet 47 (7), 702–709. https://doi.org/10.1038/ng.3285.
Rubio-Aparicio, M., Sanchez-Meca, J., Marin-Martinez, F., Lopez-Lopez, J., 2018. Guidelines for Reporting Systematic Reviews and Meta-analyses. Anales De Psicologia 34 (2), 412–420. https://doi.org/10.6018/analesps.34.2.320131.
Rusterholz*, T., Hamann, C., Markovic, A., Schmidt, S.J., Achermann, P., Tarokh, L., 2018. Nature and Nurture: Brain region specific inheritance of sleep neurophysiology in adolescence. The Journal of neuroscience : the official journal of the Society for Neuroscience 38 (43), 9275–9285. https://doi.org/10.1523/jneurosci.0945-18.2018.
Sanchez-Meca, J., Marin-Martinez, F., 2008. Confidence intervals for the overall effect size in random-effects meta-analysis. Psychological Methods 13 (1), 31–48. https://doi.org/10.1037/1082-989X.13.1.31.
Sánchez-Meca, J., López-López, J.A., López-Pina, J.A., 2013. Some recommended statistical analytic practices when reliability generalization studies are conducted. Br J Math Stat Psychol 66 (3), 402–425. https://doi.org/10.1111/j.2044-8317.2012.02057.x.
Sayón-Orea, C., Bes-Rastrollo, M., Carlos, S., Beunza, J.J., Basterra-Gortari, F.J., Martínez-González, M.A., 2013. Association between sleeping hours and siesta and the risk of obesity: the SUN Mediterranean Cohort. Obes Facts 6 (4), 337–347. https://doi.org/10.1159/000354746.
Schredl, M., Reinhard, I., 2011. Gender differences in nightmare frequency: a meta-

analysis. Sleep Med Rev 15 (2), 115–121. https://doi.org/10.1016/j.smrv.2010.06.002.
Sletten*, T.L., Rajaratnam, S.M.W., Wright, M.J., Zhu, G., Naismith, S., Martin, N.G., Hickie, I., 2013. Genetic and Environmental Contributions to Sleep-Wake Behavior in 12-Year-Old Twins. Sleep 36 (11), 1715–1722. https://doi.org/10.5665/sleep.3136.
Taylor*, M.J., Gregory, A.M., Freeman, D., Ronald, A., 2015. Do Sleep Disturbances and Psychotic-Like Experiences in Adolescence Share Genetic and Environmental Influences? Journal of Abnormal Psychology 124 (3), 674–684. https://doi.org/10.1037/abn0000057.
Toomey, R., Panizzon, M.S., Kremen, W.S., Franz, C.E., Lyons, M.J., 2015. A twin-study of genetic contributions to morningness-eveningness and depression. Chronobiol Int 32 (3), 303–309. https://doi.org/10.3109/07420528.2014.971366.
Tubs, A.S., Dollish, H.K., Fernandez, F., Grandner, M.A., 2019. The basics of sleep physiology and behavior. In: Grandner, M.A. (Ed.), Sleep and Health. Academic Press, London, pp. 3–10.
Viechtbauer, W., 2010. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software 36 (3), 1–48. https://doi.org/10.18637/jss.v036.i03.
Viechtbauer, W., Lopez-Lopez, J., Sanchez-Meca, J., Marin-Martinez, F., 2015. A Comparison of Procedures to Test for Moderators in Mixed-Effects Meta-Regression Models. Psychological Methods 20 (3), 360–374. https://doi.org/10.1037/met0000023.
Vink, J.M., Groot, A.S., Kerkhof, G.A., Boomsma, D.I., 2001. Genetic analysis of morningness and eveningness. Chronobiol Int 18 (5), 809–822.
Virta, J.J., Heikkilä, K., Perola, M., Koskenvuo, M., Räihä, I., Rinne, J.O., Kaprio, J., 2013. Midlife sleep characteristics associated with late life cognitive function. Sleep 36 (10), 1533–1541. https://doi.org/10.5665/sleep.3052. 1541A.
Visscher, P.M., Hill, W.G., Wray, N.R., 2008. Heritability in the genomics era–concepts and misconceptions. Nat Rev Genet 9 (4), 255–266. https://doi.org/10.1038/nrg2322.
Walch, O.J., Cochran, A., Forger, D.B., 2016. A global quantification of "normal" sleep schedules using smartphone data. Sci Adv 2 (5), e1501705. https://doi.org/10.1126/sciadv.1501705.
Watson, N.F., Buchwald, D., Harden, K.P., 2013. A Twin Study of Genetic Influences on Diurnal Preference and Risk for Alcohol Use Outcomes. Journal of Clinical Sleep Medicine 9 (12), 1333–1339. https://doi.org/10.5664/jcsm.3282.
Watson*, N.F., Buchwald, D., Vitiello, M.V., Noonan, C., Goldberg, J., 2010. A twin study of sleep duration and body mass index. J Clin Sleep Med 6 (1), 11–17.
Watson, N.F., Harden, K.P., Buchwald, D., Vitiello, M.V., Pack, A.I., Strachan, E., Goldberg, J., 2014. Sleep duration and depressive symptoms: a gene-environment interaction. Sleep 37 (2), 351–358. https://doi.org/10.5665/sleep.3412.
Watson, N.F., Harden, K.P., Buchwald, D., Vitiello, M.V., Pack, A.I., Weigle, D.S., Goldberg, J., 2012. Sleep Duration and Body Mass Index in Twins: A Gene-Environment Interaction. Sleep 35 (5), 597–603. https://doi.org/10.5665/sleep.1810.
Yang, J., Zeng, J., Goddard, M.E., Wray, N.R., Visscher, P.M., 2017. Concepts, estimation and interpretation of SNP-based heritability. Nat Genet 49 (9), 1304–1310. https://doi.org/10.1038/ng.3941.
Zhang, B., Wing, Y.K., 2006. Sex differences in insomnia: a meta-analysis. Sleep 29 (1), 85–93. https://doi.org/10.1093/sleep/29.1.85.