Sleep Duration and All-Cause Mortality: A Systematic Review and Meta-Analysis of Prospective Studies

Francesco P. Cappuccio, MD, FRCP; Lanfranco D’Elia, MD; Pasquale Strazzullo, MD; Michelle A. Miller, PhD

1University of Warwick, Warwick Medical School, Clinical Sciences Research Institute, Coventry, UK; 2Department of Clinical and Experimental Medicine, “Federico II” University of Naples Medical School, Naples, Italy

Background: Increasing evidence suggests an association between both short and long duration of habitual sleep with adverse health outcomes.

Objectives: To assess whether the population longitudinal evidence supports the presence of a relationship between duration of sleep and all-cause mortality, to investigate both short and long sleep duration and to obtain an estimate of the risk.

Methods: We performed a systematic search of publications using MEDLINE (1966-2009), EMBASE (from 1980), the Cochrane Library, and manual searches without language restrictions. We included studies if they were prospective, had follow-up >3 years, had duration of sleep at baseline, and all-cause mortality prospectively. We extracted relative risks (RR) and 95% confidence intervals (CI) and pooled them using a random effect model. We carried out sensitivity analyses and assessed heterogeneity and publication bias.

Results: Overall, the 16 studies analyzed provided 27 independent cohort samples. They included 1,382,999 male and female participants (follow-up range 4 to 25 years), and 112,566 deaths. Sleep duration was assessed by questionnaire and outcome through death certification. In the pooled analysis, short duration of sleep was associated with a greater risk of death (RR: 1.12; 95% CI 1.06 to 1.18; P < 0.01) with no evidence of publication bias (P = 0.74) but heterogeneity between studies (P = 0.02). Long duration of sleep was also associated with a greater risk of death (1.30; [1.22 to 1.38]; P < 0.0001) with no evidence of publication bias (P = 0.18) but significant heterogeneity between studies (P < 0.0001).

Conclusion: Both short and long duration of sleep are significant predictors of death in prospective population studies.

Keywords: Sleep duration, mortality, meta-analysis

Citation: Cappuccio FP; D’Elia L; Strazzullo P; Miller MA. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. SLEEP 2010;33(5):585-592.

SLEEP IS AN ANCESTRAL AND PRIMITIVE BEHAVIOR THAT IS SHARED ACROSS THE PLANET BY OVER A BILLION PEOPLE ON A DAILY BASIS. ITS underlying mechanisms, interactions, and long-term effects are still poorly understood. Sleep patterns of quantity and quality are affected by a variety of cultural, social, psychological, behavioral, pathophysiological, and environmental influences and have shown secular trends alongside changes in the modern society requiring longer hours of work, more shift-work and 24-7 availability of commodities, curtailing the duration of sleep to fewer hours per day across westernized populations.1 This has led to increased reporting of fatigue, tiredness, and excessive daytime sleepiness.2 Lack of sleep also exerts deleterious effects on a variety of systems with detectable changes in metabolic, endocrine,3,4 and immune pathways.5 Over the last few decades there has been growing evidence to suggest that too little sleep and too much sleep are associated with adverse health outcomes, including total mortality,6-10 type 2 diabetes,11 hypertension12,13 and respiratory disorders,14 obesity in both children and adults,15,16 and poor self-rated health.17 The relationship between duration of sleep and mortality has been often described as a U-shaped association, although other studies have not found such a uniform effect or have found no association. It is believed that different mechanisms may underlie such associations at either end of the distribution of sleep duration.18 Given the variety of studies, the large differences in the types and sizes of populations examined, the duration of follow-up, and the size of the effects, it is difficult to draw immediate conclusions on the consistency of the associations at either end of the distribution of sleep duration and at its effect size. The latter is important in public health to ascertain the likely impact at population level, if amenable to modification. The aims of this study were to (a) systematically review published prospective population-based studies, (b) carry out a meta-analysis to assess whether the global evidence supports the presence of a relationship between either short or long duration of sleep and all-cause mortality, and (c) obtain a quantitative estimate of the risk to assess the consistency and potential public health relevance. A meta-analysis allows the combination of the results of different studies, increasing the overall statistical power and the precision of estimates while controlling for bias and limiting random error.

METHODS

Literature search

We developed a search strategy to identify studies that reported the longitudinal association between duration of sleep and all-cause mortality. We searched the electronic databases MEDLINE (from 1966 to March 2009) and EMBASE (from 1980 to March 2009), as well as the Cochrane Library using the terms “sleep,” “mortality,” and “death.” Furthermore, we reviewed reference lists of original and review articles to search for more studies. Only those that were published as full-length articles were considered. No language restriction was applied.
Inclusion and Exclusion Criteria

For inclusion, studies had to fulfill the following criteria: (a) original article, (b) prospective cohort design, (c) assessment of duration of sleep as baseline exposure, (d) death recorded prospectively as outcome, (e) follow-up of at least 3 years, (f) adult population, (g) indication of the number of subjects exposed and of the rate or number of events in different sleep duration categories. Studies were excluded if: (a) case-control design was used. If multiple published reports from the same study were available, we included only the one with the most detailed information for both exposure and outcome.

Data Extraction

Data were extracted independently by two investigators (FPC and MAM), and differences were resolved by discussion and consensus. Relevant data included the first author’s surname, year of publication, country of origin of the population studied, recruitment year, number of participants, number of deaths in each group, participants’ age, gender, duration of follow-up, method used to measure sleep duration, reference category of sleep, category for “short” and “long” sleep, outcome assessment, relative risks or hazard ratios of death by both short and long sleep duration, corresponding 95% CI, and covariates adjusted in the statistical analysis.

Definition of “Short” and “Long” Sleep Duration

In each study we identified the reference category, being 7 h per night in the majority of studies,6,7,9,19-22 7 to 8 h,23-26 7 to 9 h,27-28 6 to 8 h,10 and 9 h.29 Short sleep was ≤ 5 h per night,6,7,20,22 ≤ 6 h,10,19,23,25 < 7 h,21,24,26-28 ≤ 4 h.9 Long sleep was defined as > 8 h per night,10,19,24,30 ≥ 9 h,6,7,20,22,23,25-28 ≥ 10 h,6,21 and ≥ 12 h.29

Statistical Analysis

The quality of the studies included in the meta-analysis was evaluated by the Downs & Black Quality Index score system,31 a validated checklist for assessing the quality of both randomized clinical trials and non-randomized studies. It consists of several items distributed between 5 subscales: reporting, external validity, bias, confounding, and power. For the assessment of non-randomized, prospective studies, the maximum score is 19. Relative risks (RR) or hazard ratios (HR) were extracted from the selected publications and were used to measure the relationship between sleep duration and death. Their standard errors were calculated from the respective confidence intervals. The value from each study and the corresponding standard error were transformed into their natural logarithms to stabilize the variances and to normalize the distribution. The pooled RR (and 95% CI) was estimated using a random effect model. By comparison with the reference category of sleep duration, we estimated the pooled risk and 95% CI of death for the short and for the long sleep category, separately. The heterogeneity among studies was tested by Q-statistic and quantified by H-statistic and I²-statistic.32 Funnel plot asymmetry was used to detect publication bias, and Egger’s regression test was applied to measure funnel plot asymmetry.33 We recalculated the combined estimate after estimating from the asymmetry of the funnel plot the number of “missing” studies and their effect sizes and standard errors, a method known as “trim and fill.” The influence of individual studies, from which the meta-analysis estimates are derived, was examined by omitting one study at a time to see the extent to which inferences depend on a particular study or group of studies (sensitivity analysis). Subgroup analysis was carried out to assess possible sources of heterogeneity and to check for the potential impact of age, gender, socioeconomic status, definition of sleep duration, duration of follow-up, and geographic location on the relationship between duration of sleep and all-cause mortality. All statistical analyses were performed using MIX software version 1.7.34

Role of the Funding Source

The study is part of the Sleep, Health & Society Programme of The University of Warwick. We did not receive funding for this project. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

Characteristics of Study Cohorts

Based on our selection criteria, 17 studies were deemed suitable (Figure 1). One study was reported in two separate publications,9,32 and only one was included in the analysis. Where results were reported for men and women separately, they were entered into the analyses as separate cohorts. Therefore 16 studies reporting on 27 cohorts were included in the meta-analysis. We were unable to include results from the Alameda County Study due to lack of data, although 3 separate publications had been identified.36-38 Table 1 summarizes the characteristics of the included studies. Overall, the meta-analysis included 1,382,999 participants from 8 different countries (4 studies from the USA, 4 from Japan, 3 from the UK, and 1 each from Spain, Sweden, Finland, Israel, and Taiwan). Fourteen studies recruited both men and women, while 2 studies recruited only women. Follow-up ranged from 4 to 25 years. All studies assessed death through death certificates. Sleep duration was assessed by questionnaire in all studies. The total number of deaths reported was 112,566. Of the 14 studies that included both men and women, 10 reported outcomes separately for men and women.7,9,19,21,24-28,30 Overall, there were 25 cohorts available for the relationship between short sleep and mortality and
<table>
<thead>
<tr>
<th>Author</th>
<th>Year, Publication</th>
<th>Country</th>
<th>Cohort</th>
<th>Year, Baseline</th>
<th>Gender</th>
<th>Sample size</th>
<th>Deaths</th>
<th>Follow-up, y</th>
<th>Age, y (range)</th>
<th>Quality Score</th>
<th>Exposure assessment</th>
<th>Outcome assessment</th>
<th>Adjusted variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsubono</td>
<td>1993</td>
<td>Japan</td>
<td>Mihaig Prefecture</td>
<td>1988</td>
<td>Men &amp; Women</td>
<td>4,318</td>
<td>207</td>
<td>4.0</td>
<td>61.4 (40+)</td>
<td>16</td>
<td>Questionnaire</td>
<td>Town's Death Register</td>
<td>age, sex</td>
</tr>
<tr>
<td>Ruigomez</td>
<td>1995</td>
<td>Spain</td>
<td>Health Interview Survey of Barcelona</td>
<td>1986</td>
<td>Men</td>
<td>395</td>
<td>594</td>
<td>224</td>
<td>74.3 (65+)</td>
<td>15</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, education, perceived health status</td>
</tr>
<tr>
<td>Gale</td>
<td>1998</td>
<td>UK</td>
<td>DHSS</td>
<td>1973-4</td>
<td>Men &amp; Women</td>
<td>1,209</td>
<td>1,158</td>
<td>23.0</td>
<td>75.4 (65+)</td>
<td>15</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, BMI, systolic BP, social class, diagnosis</td>
</tr>
<tr>
<td>Kojima</td>
<td>2000</td>
<td>Japan</td>
<td>Shirakawa Town</td>
<td>1982-6</td>
<td>Men</td>
<td>2,438</td>
<td>2,884</td>
<td>147</td>
<td>46.9 (20-67)</td>
<td>15</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, history of hypertension, stroke, heart and renal diseases and diabetes, use of sleeping pills, smoking and alcohol intake</td>
</tr>
<tr>
<td>Heslop</td>
<td>2002</td>
<td>UK</td>
<td>Scottish workplaces</td>
<td>1970-3</td>
<td>Men</td>
<td>6,022</td>
<td>1,006</td>
<td>2,303</td>
<td>25.0</td>
<td>(±65)</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, marital status, social class, risk factors, self-perceived stress</td>
</tr>
<tr>
<td>Kripke</td>
<td>2002</td>
<td>USA</td>
<td>CPS I</td>
<td>1982</td>
<td>Men</td>
<td>480,841</td>
<td>636,095</td>
<td>45,200</td>
<td>58.0 (30-102)</td>
<td>16</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, race, BMI, education, occupation, marital status, exercise, smoking, chuchonging, intake of fat, and fiber, insomnia, History of heart disease, hypertension, cancer, stroke, diabetes, stroke, respiratory disease, kidney disease, medications, including sleeping pills</td>
</tr>
<tr>
<td>Mallon</td>
<td>2002</td>
<td>Sweden</td>
<td>County of Dalarna</td>
<td>1983</td>
<td>Women</td>
<td>906</td>
<td>946</td>
<td>165</td>
<td>56.0 (45-65)</td>
<td>18</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, social class, country of origin, education, health status, PH of diabetes, CHD, stroke, CHF, smoking, alcohol consumption, physical activity, blood pressure, BMI, gluctose, thioycanate, creatitine, albumin, TC, HDL, homocysteine, iaesta.</td>
</tr>
<tr>
<td>Bursazeri</td>
<td>2003</td>
<td>Israel</td>
<td>Kiryat Yovel Community Health Study</td>
<td>1985-7</td>
<td>Men</td>
<td>760</td>
<td>915</td>
<td>198</td>
<td>64.0 (50+)</td>
<td>17</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, SBP, total cholesterol, BMI, smoking, alcohol consumption, education, marital status</td>
</tr>
<tr>
<td>Amagai</td>
<td>2004</td>
<td>Japan</td>
<td>Jichi Medical School Cohort Study</td>
<td>1992-6</td>
<td>Men</td>
<td>4,419</td>
<td>6,006</td>
<td>289</td>
<td>55.0 (40-69)</td>
<td>15</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, smoking, alcohol consumption, physical activity, depression, snoring, BMI, history of cancer, CVD, hypertension, diabetes, and shift work</td>
</tr>
<tr>
<td>Patel</td>
<td>2004</td>
<td>USA</td>
<td>Nurses' Health Study</td>
<td>1986</td>
<td>Women</td>
<td>82,969</td>
<td>5,409</td>
<td>14.0</td>
<td>53.4 (30-55)</td>
<td>17</td>
<td>Questionnaire</td>
<td>Death from population register</td>
<td>age, smoking, alcohol consumption, physical activity, depression, snoring, BMI, history of cancer, CVD, hypertension, diabetes, and shift work</td>
</tr>
<tr>
<td>Ferrie</td>
<td>2007</td>
<td>UK</td>
<td>Whitehall II Study</td>
<td>1985-8</td>
<td>Men &amp; Women</td>
<td>9,871</td>
<td>566</td>
<td>17.1</td>
<td>44.6 (35-55)</td>
<td>18</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, sex, marital status, employment grade, smoking, physical activity, alcohol consumption, self-rated health, BMI, SBP, total cholesterol, physical illness, GHQ, prevalent CHD</td>
</tr>
<tr>
<td>Hublin</td>
<td>2007</td>
<td>Finland</td>
<td>Finnish Twin Cohort</td>
<td>1975-81</td>
<td>Men</td>
<td>9,529</td>
<td>10,265</td>
<td>1,850</td>
<td>40.7</td>
<td>17</td>
<td>Questionnaire</td>
<td>Death from population register</td>
<td>age, education, marital status, working status, social class, BMI, smoking, binge drinking, alcohol intake, physical activity, and life satisfaction</td>
</tr>
<tr>
<td>Lan</td>
<td>2007</td>
<td>Taiwan</td>
<td>Survey of Health &amp; Living Status of the Elderly</td>
<td>1993</td>
<td>Men</td>
<td>1,748</td>
<td>1,331</td>
<td>816</td>
<td>71.3 (64+)</td>
<td>17</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, marital status, income, smoking, alcohol, BMI, exercise, depression, PMH of heart disease, stroke and cancer, naps</td>
</tr>
<tr>
<td>Gangwisch</td>
<td>2008</td>
<td>USA</td>
<td>NHANES I</td>
<td>1982-4</td>
<td>Men &amp; Women</td>
<td>5,806</td>
<td>3,903</td>
<td>273</td>
<td>45.0 (32-59)</td>
<td>18</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, sex, physical activity, smoking, depression, education, living alone, low income, daytime sleepiness, nighttime awakening, ethnicity, sleeping pill use, weight, hypertension, diabetes, general health, cancer</td>
</tr>
<tr>
<td>Ikehara</td>
<td>2009</td>
<td>Japan</td>
<td>JACC Study</td>
<td>1988-90</td>
<td>Men</td>
<td>41,489</td>
<td>57,145</td>
<td>8,548</td>
<td>58.8 (40-79)</td>
<td>18</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, BMI, PH of hypertension, diabetes, alcohol, smoking, education, exercise, employment, mental stress, depression, fresh fish intake</td>
</tr>
<tr>
<td>Stone</td>
<td>2009</td>
<td>USA</td>
<td>SOF Study</td>
<td>1993-4</td>
<td>Women</td>
<td>8,101</td>
<td>1,922</td>
<td>6.9</td>
<td>77.0 (69+)</td>
<td>18</td>
<td>Questionnaire</td>
<td>Death certificate</td>
<td>age, BMI,PH of diabetes, Parkinson's, dementia, COPD, non-skin cancer, osteoarthritis, CVD, hypertension, walks, alcohol use, smoking, depression, cognitive status, estrogen, and hypnotic use</td>
</tr>
</tbody>
</table>

1Ref. 31
Short Duration of Sleep and Mortality

Data on the relationship between short sleep and mortality are shown in Figure 2. In the pooled analysis, short duration of sleep (20 cohorts from 15 studies, n = 1,381,324 with 112,163 deaths) was associated with a greater risk of death (RR: 1.12; 95% CI 1.06 to 1.18; P < 0.01), with no evidence of publication bias (P = 0.74; Figure 3a), and no missing study identified by the trim and fill method, but significant heterogeneity between studies (I² = 39%, P = 0.02).

Long Duration of Sleep and Mortality

Data on the relationship between long sleep and mortality are shown in Figure 4. Long duration of sleep (27 cohorts available for the relationship between long sleep and all-cause mortality in population-based studies around the world. It adds to a recent meta-analysis published after our original submission.39 Our study shows an unambiguous and consistent pattern of increased risk of dying on either end of the distribution of sleep duration. Pooled analyses indicate that short sleepers (commonly < 7 h per night, often < 5 h per night) have a 12% greater risk, and long sleepers (commonly > 8 or 9 h per night) a 30% greater risk of dying than those sleeping 7 to 8 h per night. These results confirm the presence of a U-shaped association, with some heterogeneity among studies, no presence of publication bias, and a high statistical power

### Long Duration of Sleep and Mortality

The effect was consistent in younger (< 60 years) and older (≥ 60 years) cohorts, in men and women, whether or not socioeconomic status was adjusted for (Figure 5a). The effect did not differ by variations in the definition of short sleep, length of follow-up, or geographic location (Figure 5a).

### DISCUSSION

This study provides a comprehensive and systematic review of the literature and quantitative estimates of the longitudinal associations between habitual short and long duration of sleep and all-cause mortality in population-based studies around the world. It adds to a recent meta-analysis published after our original submission.39 Our study shows an unambiguous and consistent pattern of increased risk of dying on either end of the distribution of sleep duration. Pooled analyses indicate that short sleepers (commonly < 7 h per night, often < 5 h per night) have a 12% greater risk, and long sleepers (commonly > 8 or 9 h per night) a 30% greater risk of dying than those sleeping 7 to 8 h per night. These results confirm the presence of a U-shaped association, with some heterogeneity among studies, no presence of publication bias, and a high statistical power
conferring by over 1.3 million participants included with more than 110,000 events. The studies allowed for a large number of potential confounders. Furthermore, for short sleep the effects were comparable between men and women, did not depend on age, socioeconomic status, definition of short duration of sleep (≤ 5, ≤ 6 h, or ≤ 7 h per night), duration of follow-up, or geographic location. In contrast, for long sleep, the estimate of effect was stronger in older cohorts, in studies where long sleep was defined as ≥ 10 h per night, in follow-ups not longer than 19 years, and in East Asian cohorts.

These results are of interest for several reasons. First, the direction of the associations is consistent in different populations. Although the meta-analysis detected significant heterogeneity between studies, further sensitivity analysis and the absence of publication bias are in favor of consistent effects across populations. Second, they indicate an effect size of potential relevance. Third, the consistency of assessments of both exposure (duration of sleep by questionnaire) and outcomes (death certificates) across studies limits the degree of variability due to differences in methods.

**Study Limitations**

First, the quality of the data cannot go beyond the quality of the individual studies included. Second, a meta-analysis of observational data, though prospective, is open to important fallacies in that it cannot directly control for confounding. We made an attempt to allow for multiple confounding by including adjusted estimates from multivariate models from each contributing study. However, residual confounding and bias

![Figure 3](image_url)

**Figure 3**—Funnel plots for meta-analyses of short (a) and long (b) duration of sleep and mortality. Egger’s test: P = 0.74 for short sleep and P = 0.18 for long sleep.

<table>
<thead>
<tr>
<th>First author Year</th>
<th>Sample size</th>
<th>Deaths</th>
<th>Long sleep v Reference</th>
<th>Relative Risk 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsubono 1993</td>
<td>4,318</td>
<td>207</td>
<td>≥9 v 7-8h</td>
<td>1.58 (1.16 to 2.17)</td>
</tr>
<tr>
<td>Ruigomez (men) 1995</td>
<td>395</td>
<td>224</td>
<td>&gt;9 v 7-9h</td>
<td>1.30 (0.71 to 2.38)</td>
</tr>
<tr>
<td>Ruigomez (women) 1995</td>
<td>594</td>
<td></td>
<td>&gt;9 v 7-9h</td>
<td>1.46 (0.80 to 2.68)</td>
</tr>
<tr>
<td>Gale 1996</td>
<td>1,299</td>
<td>1,158</td>
<td>≥12 v 9h</td>
<td>1.70 (1.17 to 2.47)</td>
</tr>
<tr>
<td>Kojima (men) 2000</td>
<td>2,438</td>
<td>147</td>
<td>9-9.9 v 7-8.9h</td>
<td>1.15 (0.75 to 1.77)</td>
</tr>
<tr>
<td>Kojima (women) 2000</td>
<td>2,884</td>
<td>109</td>
<td>&gt;8 v 7-8h</td>
<td>1.07 (0.58 to 1.97)</td>
</tr>
<tr>
<td>Heslop (men) 2002</td>
<td>6,022</td>
<td>2,303</td>
<td>&gt;8 v 7-8h</td>
<td>0.81 (0.68 to 0.97)</td>
</tr>
<tr>
<td>Heslop (women) 2002</td>
<td>1,006</td>
<td>262</td>
<td>≥9 v 7h</td>
<td>1.20 (0.71 to 2.03)</td>
</tr>
<tr>
<td>Kripke (men) 2002</td>
<td>480,841</td>
<td>45,200</td>
<td>9 v 7h</td>
<td>1.17 (1.13 to 1.22)</td>
</tr>
<tr>
<td>Kripke (women) 2002</td>
<td>636,095</td>
<td>32,440</td>
<td>&gt;8 v 7h</td>
<td>1.23 (1.19 to 1.28)</td>
</tr>
<tr>
<td>Mallon (men) 2002</td>
<td>906</td>
<td>165</td>
<td>&gt;8 v 7h</td>
<td>1.99 (1.22 to 3.25)</td>
</tr>
<tr>
<td>Mallon (women) 2002</td>
<td>964</td>
<td>101</td>
<td>&gt;8 v 7h</td>
<td>1.30 (0.83 to 2.08)</td>
</tr>
<tr>
<td>Burazeri (men) 2003</td>
<td>760</td>
<td>198</td>
<td>&gt;8 v &lt;6h</td>
<td>2.14 (1.24 to 3.70)</td>
</tr>
<tr>
<td>Burazeri (women) 2003</td>
<td>915</td>
<td>205</td>
<td>≥9 v 7-7.9h</td>
<td>0.80 (0.51 to 1.26)</td>
</tr>
<tr>
<td>Amagai (men) 2004</td>
<td>4,419</td>
<td>289</td>
<td>≥9 v 7-7.9h</td>
<td>1.11 (0.78 to 1.57)</td>
</tr>
<tr>
<td>Amagai (women) 2004</td>
<td>6,906</td>
<td>206</td>
<td>≥9 v 7-7.9h</td>
<td>1.51 (0.98 to 2.32)</td>
</tr>
<tr>
<td>Patel 2004</td>
<td>82,969</td>
<td>5,409</td>
<td>≥9 v 7h</td>
<td>1.40 (1.27 to 1.55)</td>
</tr>
<tr>
<td>Ferrie 2007</td>
<td>9,871</td>
<td>566</td>
<td>≥9 v 7h</td>
<td>1.25 (0.93 to 1.67)</td>
</tr>
<tr>
<td>Hublin (men) 2007</td>
<td>9,529</td>
<td>1,850</td>
<td>≥9 v 7-8h</td>
<td>1.25 (1.09 to 1.45)</td>
</tr>
<tr>
<td>Hublin (women) 2007</td>
<td>10,265</td>
<td>1,850</td>
<td>≥9 v 7-8h</td>
<td>1.17 (1.02 to 1.35)</td>
</tr>
<tr>
<td>Lan (men) 2007</td>
<td>1,748</td>
<td>816</td>
<td>≥10 v 7-7.9h</td>
<td>1.51 (1.19 to 1.91)</td>
</tr>
<tr>
<td>Lan (women) 2007</td>
<td>1,331</td>
<td>522</td>
<td>≥10 v 7-7.9h</td>
<td>2.05 (1.50 to 2.81)</td>
</tr>
<tr>
<td>Gangwisch (32-59) 2008</td>
<td>5,806</td>
<td>273</td>
<td>≥9 v 7h</td>
<td>1.04 (0.66 to 1.63)</td>
</tr>
<tr>
<td>Gangwisch (50-86) 2008</td>
<td>3,983</td>
<td>1,604</td>
<td>≥9 v 7h</td>
<td>1.36 (1.17 to 1.59)</td>
</tr>
<tr>
<td>Ikehara (men) 2009</td>
<td>41,489</td>
<td>8,546</td>
<td>≥10 v 7h</td>
<td>1.40 (1.27 to 1.55)</td>
</tr>
<tr>
<td>Ikehara (women) 2009</td>
<td>57,145</td>
<td>5,992</td>
<td>≥10 v 7h</td>
<td>1.55 (1.38 to 1.75)</td>
</tr>
<tr>
<td>Stone 2009</td>
<td>8,101</td>
<td>1,922</td>
<td>&gt;8 v 6-8h</td>
<td>1.16 (0.97 to 1.39)</td>
</tr>
<tr>
<td>Combined effect: P &lt; 0.0001</td>
<td>1,382,999</td>
<td>112,566</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: I^2 = 71%; Q = 90.1, P < 0.0001

Publication bias: Egger’s test: P = 0.18

**Figure 4**—Forest plot of the risk of death associated with long duration of sleep compared to the reference group in 27 population cohorts from 16 published prospective studies including 1,382,999 participants and 112,566 events. Results are expressed as relative risk (RR) and 95% confidence intervals (95% CI). Pooled analysis P < 0.0001; heterogeneity test: I^2 = 71%, P < 0.0001
remain a possibility. Third, the results can only be representative of the studies that have been included and are unable to provide a representative inference of all studies published, but not included. Nevertheless, there was no evidence of publication bias. These results are therefore important in guiding the assessment of current evidence and the definition of future research strategies and public health policy decisions.

All studies used sleep questionnaires to determine self-reported sleep duration within their populations. This method often did not allow (unless explicitly built as additional questions) to differentiate time asleep from time in bed or to estimate number and duration of naps. It is not usually feasible to obtain more detailed and objective measures of sleep in large prospective population studies. Sleep diaries, actigraphy, and polysomnography from some large population and small-scale investigations have shown high correlations between subjective estimates of sleep duration and the more direct assessments. However, in the few examples in the literature their use confirmed the results obtained previously with self-reported sleep assessment. Furthermore, assessments of sleep duration in the primary health care setting rely on self-reported data from patients.

Sleep duration was assessed at one point in time in all studies, and it is possible that a single measure of exposure may not fully capture the sustained effects of sleep duration over time when relating them to long-term mortality. Changes in sleep duration over time may represent a better measure of exposure in this context. Three studies have addressed this issue by measuring change in sleep duration over time as a proxy for prolonged exposure to short or long sleep duration in relation to mortality. In the West of Scotland study, participants were assessed twice 4 to 7 years apart and followed up for 25 years. Those who reported at both time points < 7 h of sleep showed an increased risk of death from all causes (men 1.15 [0.93 to 1.42] and women 1.73 [0.99 to 3.03]), as well as an increased risk of death from cardiovascular causes (men 1.19 [0.90 to 1.58] and women 2.30 [0.94 to 5.60]). In the Whitehall II Study, changes in sleep duration were assessed 5 years apart, and follow-up lasted nearly 12 years. Both a decrease and an increase in sleep duration over time were associated with increased all-cause mortality (1.62 [1.17 to 2.25] and 1.75 [1.24 to 2.47], respectively). However, while the increased risk of death in those who had increased their sleep was mainly attributable to an increase in non-cardiovascular deaths (2.06 [1.38 to 3.08]), those who had curtailed their duration of sleep had an increased risk of death predominantly from cardiovascular causes (2.04 [1.20 to 3.49]). In a large Finnish study of twins, changes in duration of sleep over 6 years (either to shorter or longer sleep) were associated to increased mortality in both men and women. Finally, it is likely that the study samples included in the analysis did not exclude subjects with obstructive sleep apnea-hypopnea syndrome (OSAS). These would represent approximately 4% of middle-aged men and 2% of middle-aged women. OSAS is associated with obesity, disrupted short sleep, excessive daytime sleepiness and high rates of morbidity and mortality, predominantly due to cardiovascular disease. Whilst it is possible that they may have contributed to the observed increased risk of death amongst short sleepers, the adjustment for obesity or body mass index in almost every study would have corrected for this.

Figure 5—Subgroup analyses to explore sources of heterogeneity in short (a) and long (b) duration of sleep.
Heterogeneity

Our results are consistent with other epidemiological evidence indicating increased risk of cardiovascular morbidity and mortality,\textsuperscript{8,9,46} coronary artery calcifications,\textsuperscript{47} cardiovascular risk predictors like hypertension,\textsuperscript{12,13,48} obesity,\textsuperscript{15,16,49,51} type 2 diabetes or impaired glucose control,\textsuperscript{11} and atherogenic lipid profile\textsuperscript{72} with shorter duration of sleep.

Gender-differences in outcomes and risk related to duration of sleep have also been reported.\textsuperscript{5,13,53-55} Our analysis was repeated after stratification by gender. No differences were detected in the association with either short or long duration of sleep and all-cause mortality. Ideally long follow up durations would be appropriate to assess the influence of sleep duration on health over the life course.\textsuperscript{22} We excluded a priori short follow-up studies (< 3 years) to avoid measurements of sleep duration being too close to the time of death and included studies with follow-up ranging from 4 to 25 years. Furthermore, a stratified analysis by duration of follow-up was carried out, that did not suggest any trend. There was some indication that the mean effects for both short and long duration of sleep were stronger in studies carried out in East Asia, predominantly Japan. The explanations may be different. For short sleep, some cross-cultural comparisons have shown that the average duration of sleep is shorter in countries of East Asia,\textsuperscript{17} possibly as a result of societal pressures. For long sleep, the stronger association with mortality may, at least in part, be the result of longer life expectancy in countries like Japan compared to Europe and the US. It is likely that the shape of the relationship between sleep duration and mortality is heavily influenced by deaths in elderly participants. This view is supported by a stronger effect of long duration of sleep on all-cause mortality in older cohorts (≥ 60 years). Finally, socioeconomic status does not appear to modify the risk estimates when we compared cohorts that did or did not adjust for this variable.

Potential Mechanisms

The mechanisms that underlie these associations are not fully understood. Causative mechanisms relating short duration of sleep to adverse health outcomes include reciprocal changes in circulating levels of leptin and ghrelin,\textsuperscript{56,57} that in turn would increase appetite, caloric intake, reduce energy expenditure,\textsuperscript{4} and facilitate the development of obesity\textsuperscript{57} and impaired glycemic control\textsuperscript{68} with increased cardiovascular risk. Increased cortisol secretion and altered growth hormone metabolism have also been implicated.\textsuperscript{69} Finally, low grade inflammation is activated during short sleep, with possible implications not only for cardiovascular disease but also for other chronic conditions including cancer.\textsuperscript{3} Conversely, no studies published to date have demonstrated a possible mechanism mediating the effect of long duration of sleep as a cause of morbidity and mortality. The association between long duration of sleep and mortality may be explained by residual confounding and co-morbidities.\textsuperscript{80} In particular, depressive symptoms, low socioeconomic status, unemployment, low level of physical activity, undiagnosed health conditions, poor general health, and cancer-related fatigue have all been shown to be associated with long duration of sleep and to confound the association with morbidity and mortality.\textsuperscript{80}

Implications

Currently, there is no evidence that sleeping habitually between 6 and 8h per day in an adult is associated with harm and long term health consequences. In terms of prevention, consistently sleeping 6 to 8h per night may therefore be optimal for health. However, sleeping 9h or more per night may represent a useful diagnostic tool for detecting subclinical or undiagnosed co-morbidity. People reporting consistently sleeping 5h or less per night should be regarded as a higher risk group for all-cause mortality. A 12% increased risk of death in short sleepers, if causally related, would equate to over 6.3 million attributable deaths in the UK in people over 16 years of age and over 25 million attributable deaths in the US in people over the age of 20 years. Future studies should be designed to answer the question whether sleep duration is a cause or simply a marker of ill-health. Indirect evidence seems to suggest to date that sustained sleep deprivation may trigger biological mechanisms contributing to the deterioration of health status, whereas long duration of sleep may be a powerful additional marker of ill-health.

Duration of sleep should be regarded as an additional behavioral risk factor or risk marker heavily determined by the environment and possibly amenable to modification through both education and counselling as well as through measures of public health. The latter would aim at favorable modifications of physical and working environments to allow sufficient sleep and avoid habitual and sustained sleep deprivation.

ACKNOWLEDGMENTS

We thank Mrs. Patricia McCabe for helping with the systematic searches. This study was supported, in part, by an EC Grant (FP7-HEALTH- 2007-201550).

Authors’ contribution: FPC conceived the study aims and design, contributed to the systematic review and data extraction, performed the analysis, interpreted the results, and drafted the manuscript. LD, PS, and MAM contributed to the data extraction, interpretation of results, and to the revision of the manuscript.

DISCLOSURE STATEMENT

This was not an industry supported study. The authors have indicated no financial conflicts of interest.

REFERENCES