Study Sample Characteristics in the Sleep and Chronobiology Research: A Systematic Review

true

true

true
Affiliation

any aspects of sleep and circadian physiology appear to be sensitive to characteristics of the studied population, most notably sex. While recent research robustly highlights the importance of considering participant-level demographic information, it is not clear to what extent this information is available in the large body of already published literature. In this systematic review, we evaluated the study sample characteristics in the published sleep and chronobiology research over the past 40 years.

Articles published between 1979 and 2019 (odd years) in the top eight sleep and chronobiology journals, identified by their five-year Impact Factor, were found through MEDLINE. 6,777 articles were initially included for screening. Inclusion requirements included conducting original research, reporting human data, and recruiting volunteers. The reporting of sample size, age, sex, gender, ethnicity, level of education, socioeconomic status, and profession of the study population was scored binarily (0=not reported), and any reported aggregate summary statistics for these variables were recorded. Funding source, geographical location and clinical focus of the article were examined, as well as whether data were analyzed including any of the demographic variables as covariates.

~75% of screened articles met inclusion criteria. While >90% of studies reported age or sex, all other variables were reported in <10% of cases. We found that sex balance greatly changed over the years, from a ~3:1 male to female ratio in the 1990s to a near-equal representation in the 2010s. Overall, ~75% of studies recruited both male and female participants. Of studies recruiting a single sex, ~50% all-female studies focused on a sex-dependent feature, compared to <5% in all-male studies.

In this comprehensive review, we found that the majority of studies report at least sex or age, while many other important variables are typically not reported. Reporting quality is highly variable, indicating an opportunity to standardize reporting guidelines for participant-level characteristics to facilitate disaggregated data analyses.

Introduction

Many aspects of sleep and circadian physiology appear to be sensitive to characteristics of the studied population, most notably sex. While recent research robustly highlights the importance of considering participant-level demographic information, it is not clear to what extent this information is available in the large body of already published literature. In this systematic review, we evaluated the study sample characteristics in the published sleep and chronobiology research over the past 40 years.

Methods

Journal articles published between 1979 and 2019 (odd years) in the top eight sleep and chronobiology journals were found through MEDLINE by MS. The eight journals were selected based on their five-year Impact Factor, and included Journal of Pineal Research, Sleep, Journal of Sleep Research, Sleep Medicine, Journal of Clinical Medicine, Journal of Biological Rhythms, Behavioral Sleep Medicine, and Chronobiology International. 6,777 articles were initially sampled for screening. Inclusion requirements included conducting original research in the English language, reporting human data, and recruiting volunteers. As such, animal studies, bibliographies, case reports, comments, conference proceedings, editorials, guidelines, letters, retracted publications, reviews, erratums and corrigendums were excluded. All articles were reviewed for eligibility and coded by ST. The reporting of sample size, age, sex, gender, ethnicity, level of education, socio-economic status, and profession of the study population was scored binarily (0 = not reported, 1 = reported), and any reported aggregate summary statistics for these variables were recorded. Funding source, geographical location and clinical focus of the article were examined, as well as whether data were analyzed by including any of the demographic variables as covariates.

Data were coded in an Excel Spreadsheet and analyzed in R Studio (version 4.0.4). Reporting of funding, geographical location, and number of sub-studies for each article were investigated for the sample of articles that passed all eligibility criteria. The rest of the analysis was conducted on the subset of articles that reported a single study (97% of articles), i.e. a single sample population.

Results

Inclusion & Exclusion. 84% of the sampled articles met eligibility criteria during the review.

```
figure1B
```

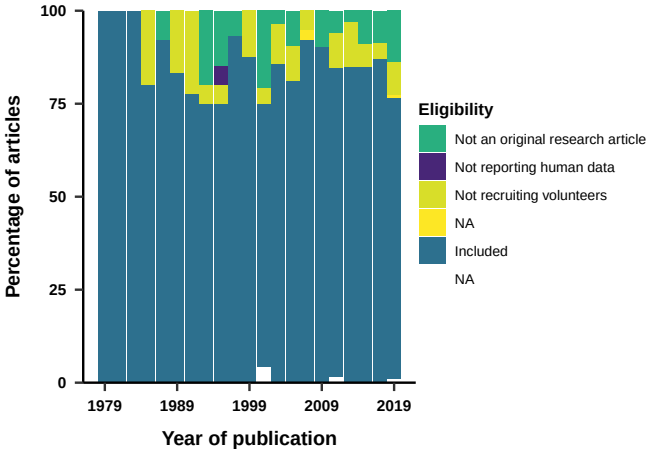


Figure 1. Inclusion and exclusion of articles.

Funding. Funding sources were reported by 61% of studies, while funding number was also reported in 70% of these cases. Overall, funding by the United States' National Institutes of Health (NIH) represented 15% of the reported funding agencies. 94% of the studies funded by the NIH also reported funding number.

Geographical location. 96% of articles were conducted in a single country. The geographical location of the study was explicitly reported in 58% of studies. The country of study was inferred for the rest of the sampled articles. Inference was primarily based on the first author's affiliation. Overall, 45 countries were represented. Figure 2B shows the distribution of study location across time with the eight most represented countries.

```
figure2B
```

Sample size. Sample size was reported in 87% of studies, while 97% investigated a single sampled population.

```
figure3
```

Age. 94% of articles reported a variable describing age. Figure 4A and 4B show the use of various age variables by year and journal of publication. Overall, the mean age of the study populations was 40 years old.

```
figure4A
```

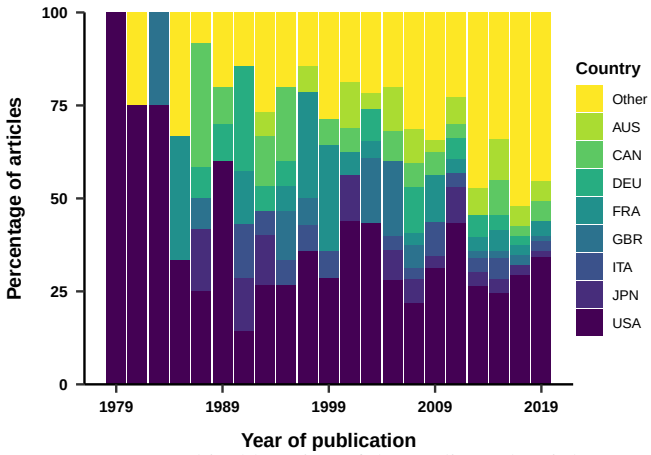


Figure 2. Geographical location of the studies. The eight most represented countries are individually shown.

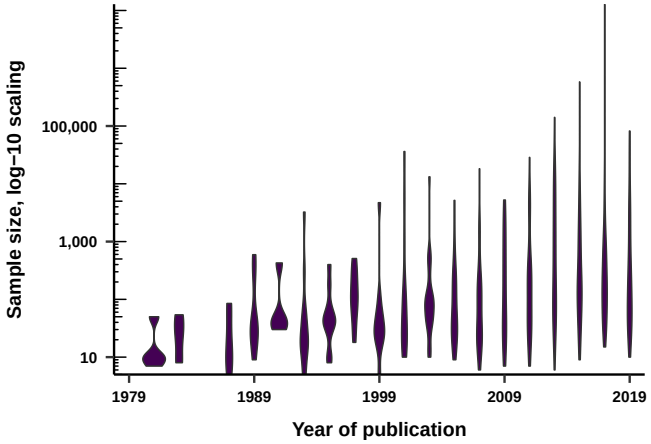
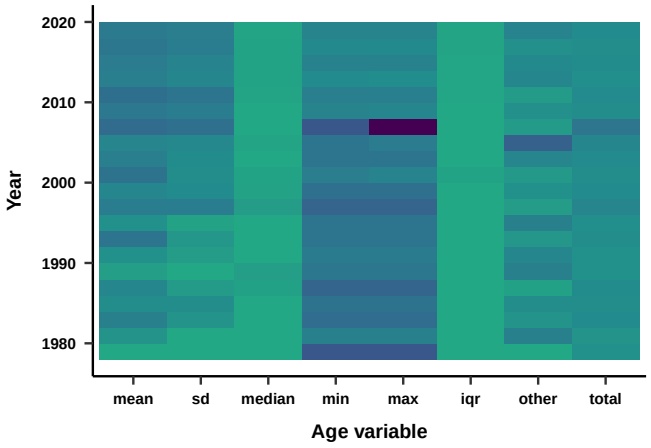
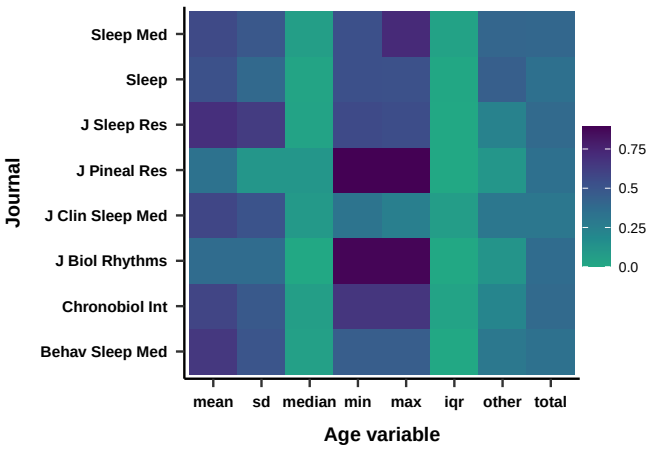


Figure 3. Sample size of the recruited volunteers. Numbers are computed with a log-10 scale.




```
figure4B
```




```
figure4C
```

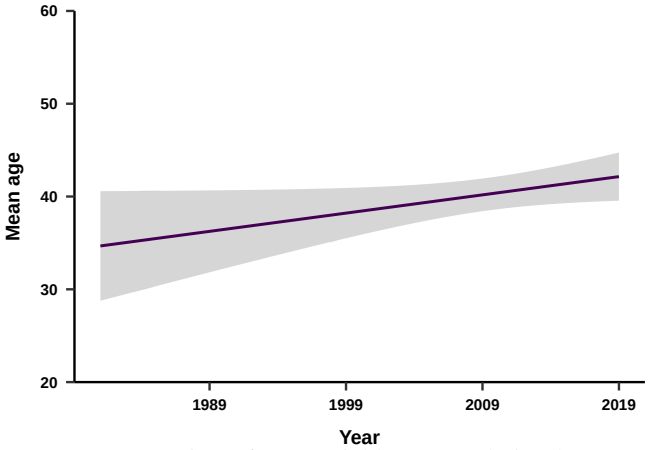


Figure 4. Reporting of age variables. Correlation between the reporting of the mean, standrad deviation of the mean, median, minimum, maximum, interquartile range and other age variables by A. year of publication and B. journal of publication. C. Mean age by year of publication.

```
#grid.arrange(figure4A, figure4C, figure4B, nrow=2)
```

Sex. Sex was reported in 93% of the studies. Figure 5 displays the proportion of studies that recruited male subjects, female subjects, both sexes or did not specify the sex of the participants. 10% of the studies reporting sex only recruited male participants, while 11% only employed females. Out of the studies focusing on a single gender, 2% of the male studies focused on a sex dependent feature, while 3% of the female studies did. 4% of studies reported age by sex.

```
figure5A
```

Ethnicity, education, profession and socio-economic status. Other demographics variables were reported in 14% of studies for education, 17% for ethnicity, 4% for profession, and 5% for socio-economic status. Figure 6 shows the distribution of this reporting across the years. Figures 7 show the number of categories reported for each variable.

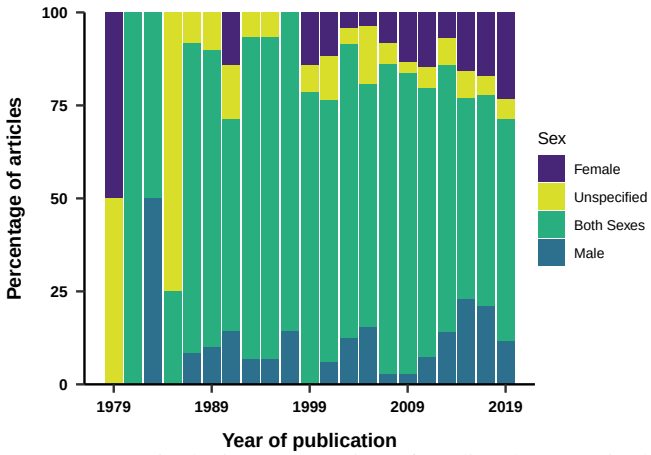


Figure 5. Sex inclusion. Proportion of studies that recruited male subjects, female subjects, both sexes, or did not specify the sex of the participants.

```
#figure6.bars
figure6.points
```

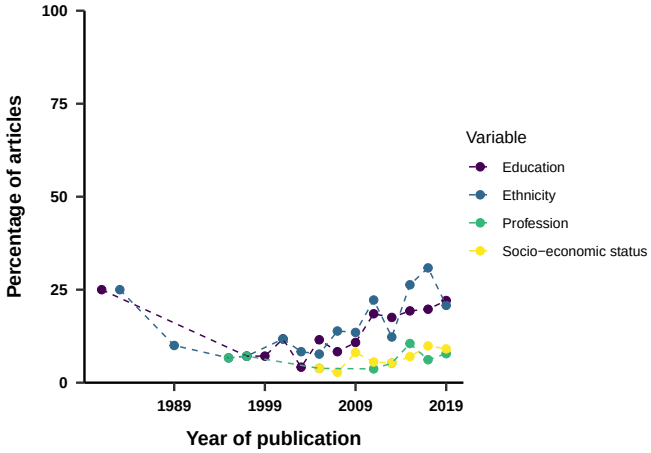


Figure 6. Reporting of education, ethinicity, profession and socio-economic status.

```
grid.arrange(figure7A, figure7B, figure7C, figure7D, nrow=
```

Study focus. 5% of articles focused on a sex dependent feature, while 49% investigated a clinical feature. 0% of studies focused on twins, 1% on pregnant women, 2% on shift workers and 5% on university students.

Analysis disaggregation. A

```
figure8
```

Discussion

Conclusion

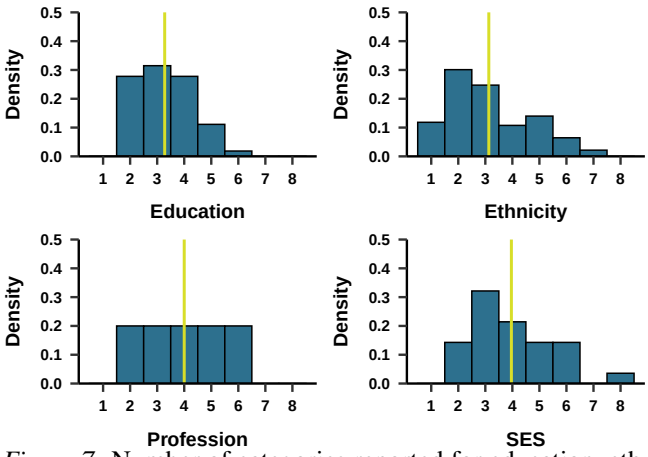


Figure 7. Number of categories reported for education, ethnicity, profession and socio-economic status.

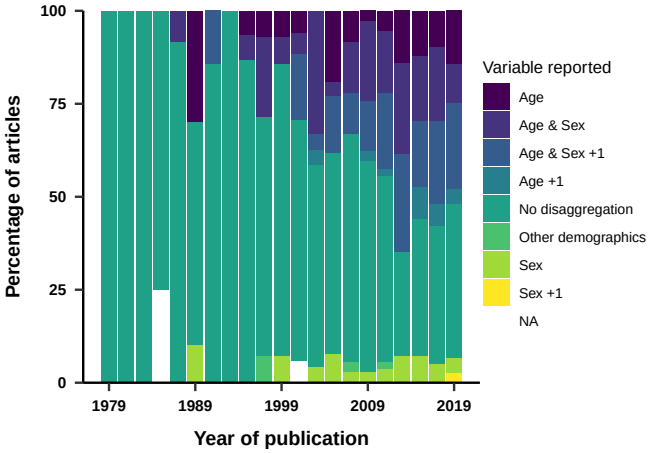


Figure 8. Use of study population characteristics as variables in the analysis.

References