Supplementary Material

Does the choice of Sleep Regularity Index software change the conclusions? A re-analysis of the NHANES 2011–2014 sleep-regularity–obesity association

Jürgen Degenfellner (ZHAW School of Health Sciences)

July 10, 2026

This supplement reports, across the three SRI calculators (the original by-hand Phillips implementation, `GGIR` and `sleepreg`), the same diagnostic displays provided in the original study, plus the verification of the package computations.

1 Verification of the package computations

On synthetic 7-day recordings with an analytically known SRI, all three calculators returned the correct value (Table S1); on a pseudo-random pattern they agreed with the hand formula. This confirms the divergence reported in the main text is a genuine consequence of algorithmic conventions, not an implementation error.

Table S1: Synthetic verification (7 consecutive days, no missing epochs).

Test case	expected	hand	`GGIR`	`sleepreg`
7 identical days	+100	100.00	100.00	100.00
alternating pattern/complement	−100	−100.00	−100.00	−100.00
agree on half the epochs	0	0.00	0.00	−0.01
shifted pseudo-random	—	5.56	5.89	5.54

Reproducibility note. `sleepreg` (v1.3.5) aborts when a recording starts exactly at local midnight (returning a missing value), and its pinned dependency `GGIR` 2.0-0 prevents installation alongside current `GGIR`; I therefore ran its SRI engine from source.

2 Exact package invocations and parameter settings

For full reproducibility I give the verbatim calls used to compute each SRI on the byte-identical binary sleep/wake series, with package versions and every argument (Table S2). The comparison uses each package's *default* settings throughout; the only non-defaults are output switches (`wr.raster`, `overwr`) that do not affect the SRI value. `sleepreg`'s SRI engine was run from its (`GGIR`-independent) source functions, verbatim from v1.3.5, because its pinned dependency `GGIR` 2.0-0 cannot coexist with current `GGIR`; the synthetic verification above confirms this reproduces the intended quantity.

```
# (1) Hand -- original Phillips formula, 1-min epochs, weekday-matched pairs
```

```
SRI = -100 + (200/N_valid) * sum( s[i] == s[i+1440] )   # over valid pairs

# (2) GGIR 3.3.6
df <- data.frame(time = as.POSIXct(ts, origin="1970-01-01", tz="UTC"),
                 sleepstate = as.numeric(s),
                 invalid    = as.integer(is.na(s)))
out <- GGIR::CalcSleepRegularityIndex(data=df, epochsize=60, desiredtz="UTC")
SRI <- mean(out$SleepRegularityIndex, na.rm=TRUE)  # mean over day-pairs

# (3) sleepreg 1.3.5 -- run from source, two-step pipeline
SWS_from_binarySW(binarySWdir=bindir, tsCol=2, binaryCol=1)
SRI_from_binary(binarydir=swsdir, col.trans=1, col.timestamp=2,
                wr.raster=FALSE, overwr=TRUE)
                # defaults kept: minSRIdays=5, exclNAhrs=6
```

Table S2: Package versions and SRI-relevant parameter settings. All SRI-relevant parameters are at package defaults; the non-default arguments (`wr.raster`, `overwr`) affect only the written output, not the SRI value.

Calculator	Version	SRI-relevant settings (package defaults)
Published (hand Phillips)	—	1-min epochs; day-of-week-matched consecutive day pairs
`GGIR`	3.3.6	`epochsize=60` (1-min epochs); consecutive calendar-day pairs, time-of-day matched; mean over pairs
`sleepreg`	1.3.5	`minSRIdays=5`; `exclNAhrs=6`; 10-s epochs; continuous 24 h lag

3 SRI distribution by calculator

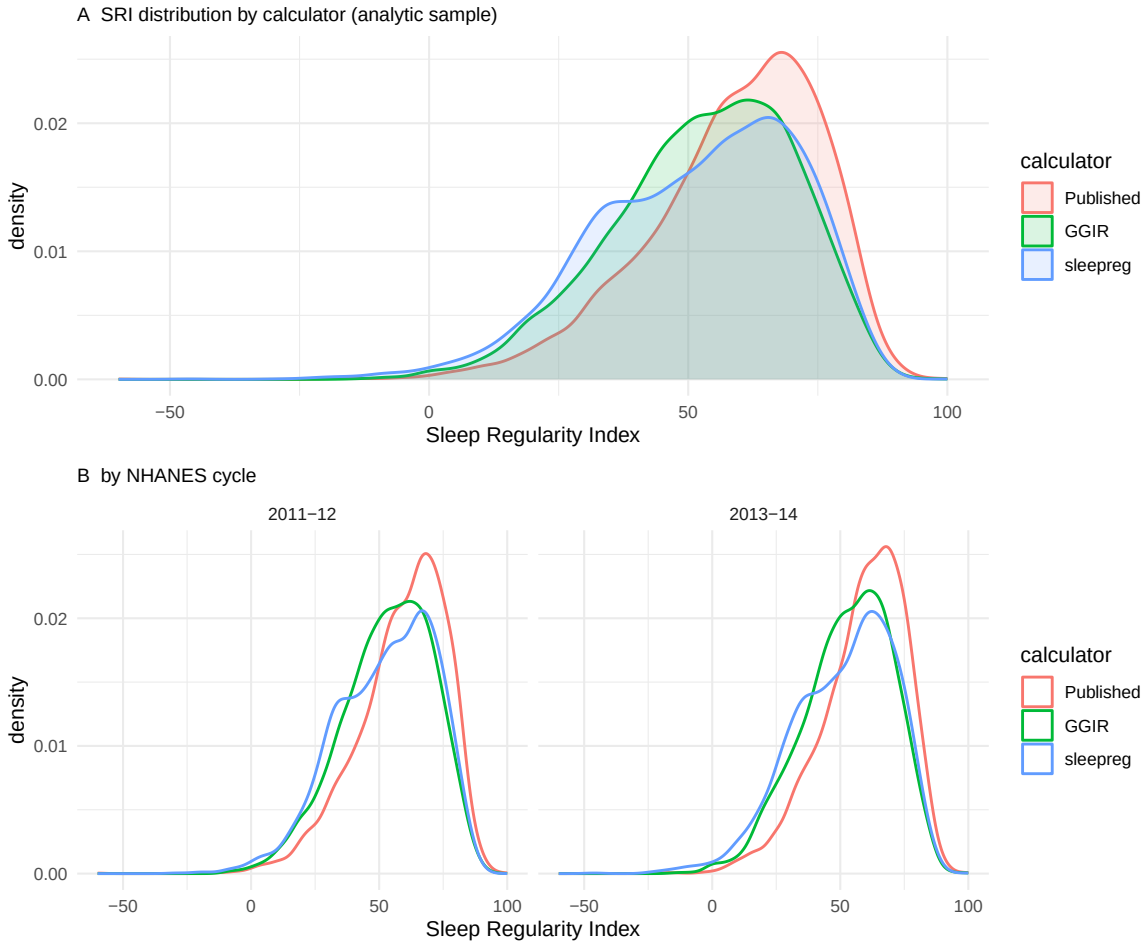

Figure S1: Distribution of the SRI by calculator on identical input: overall (A) and by NHANES cycle (B). The package distributions are shifted lower than the published SRI.

4 Pairwise agreement (Bland–Altman)

The all-versus-all pairwise agreement of the published SRI and the two packages is shown in the main text (Figure 1); here I add the complementary Bland–Altman view of each package against the published SRI.

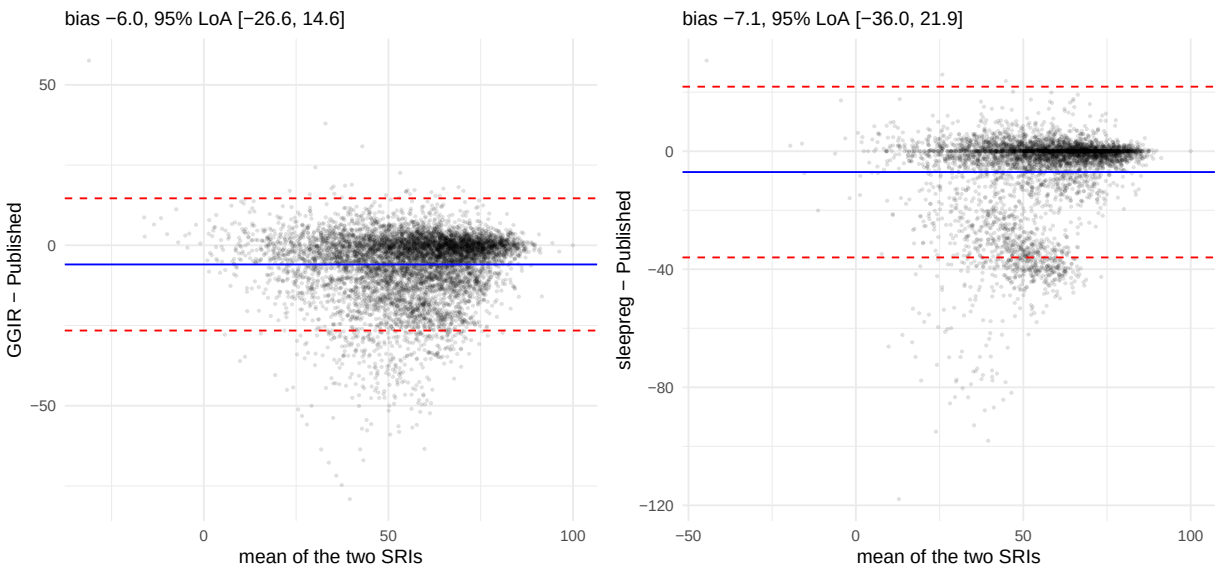

Figure S2: Bland–Altman agreement of each package SRI with the published SRI (blue: mean difference; red dashed: 95% limits of agreement).

5 Structure of the disagreement

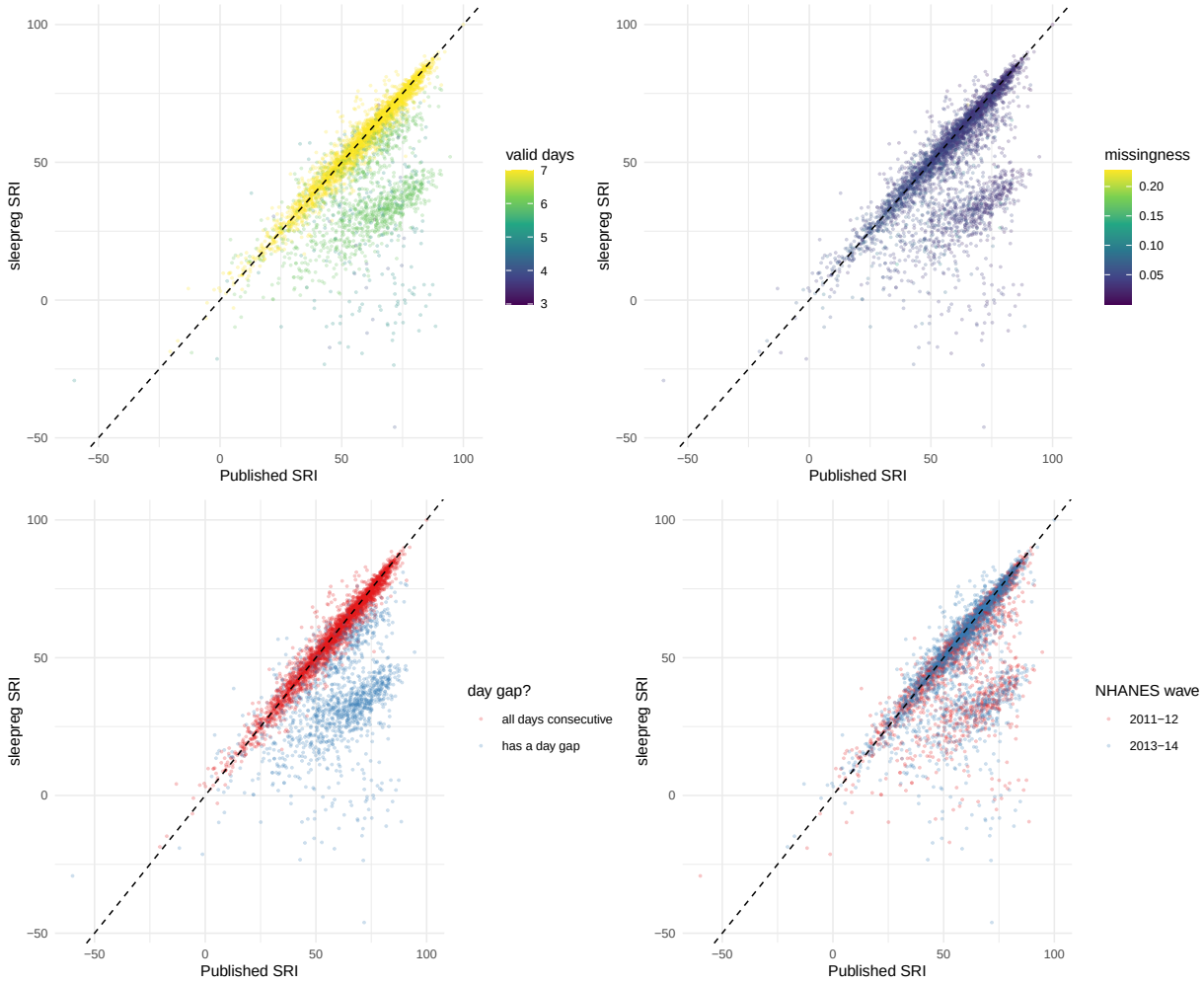

Figure S3: The `sleepreg`-versus-published scatter coloured by candidate features. The off-diagonal clouds track the number of valid (consecutive) days / the presence of a day gap, not missingness or NHANES wave.

6 The re-fitted model and predicted BMI by calculator

Throughout this section "the fully adjusted model" is the fully adjusted, non-interaction model of the original study, **Model 2 of Degenfellner et al. (2026), *Journal of Biological Rhythms* 41(2):205–218, Table 3**: log(BMI) regressed on the SRI plus age, sex, race/ethnicity, education, household income, occupation category, alcohol consumption, smoking status, vitamin D, total caloric intake, PHQ-9 depression score, marital status and physical activity, re-fitted with each SRI calculator substituted for the exposure. The residual and assumption diagnostics for this model are shown per calculator further below (`car::qqPlot()` and `performance::check_model()`).

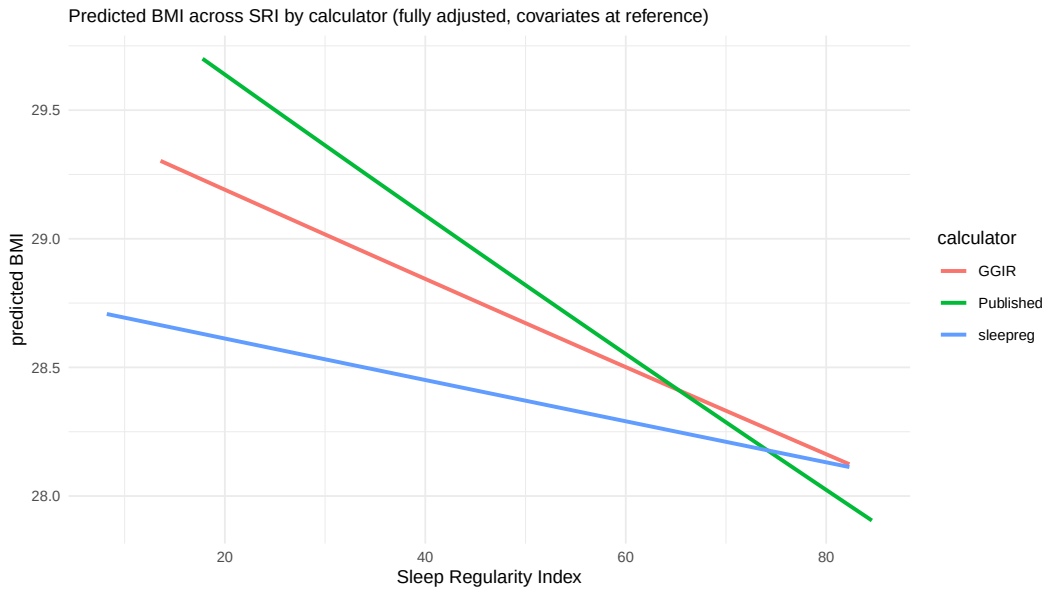

Figure S4: Predicted BMI across the SRI range under each calculator, from the fully adjusted model (Model 2 of Degenfellner et al. 2026, *J Biol Rhythms* 41(2):205–218, Table 3; covariates at reference). The slope flattens from the published SRI to `GGIR` to `sleepreg`.

7 Are the re-fitted models well specified?

I checked the assumptions of the fully adjusted model (Model 2 of Degenfellner et al. 2026, Table 3) re-fitted with each calculator (the per-calculator `car::qqPlot()` and `check_model()` panels below). The models are *equally well behaved across calculators*: adjusted $R^2 \approx 0.13$, residual skewness ≈ 0.31 and excess kurtosis ≈ 0.55, and a maximum variance-inflation factor of 2 in every case. The corresponding *magnitudes* (mild right-skew of residuals, slight heteroscedasticity, a barely-curved SRI term) are near-identical for all calculators and are already present in the original model. In other words, the attenuation of the effect under `GGIR`/`sleepreg` is not an artefact of those models fitting worse; they fit essentially as well as the published one.

8 `performance::check_model()` and `car::qqPlot()` outputs

For completeness, and matching the diagnostic displays of the original study, I reproduce the full `performance::check_model()` panel (posterior predictive check, linearity, homogeneity of variance, influential observations, collinearity, normality of residuals) for the fully adjusted model under each calculator (Figs. S5–S7), and the `car::qqPlot()` of studentized residuals with 95% envelopes (Fig. S8). The panels are visually indistinguishable across calculators: the posterior predictive and homogeneity panels are well behaved, all variance-inflation factors are below the conventional concern level, no observation falls outside the influence contours, and the QQ panels show the same mild heavy right tail in every case.

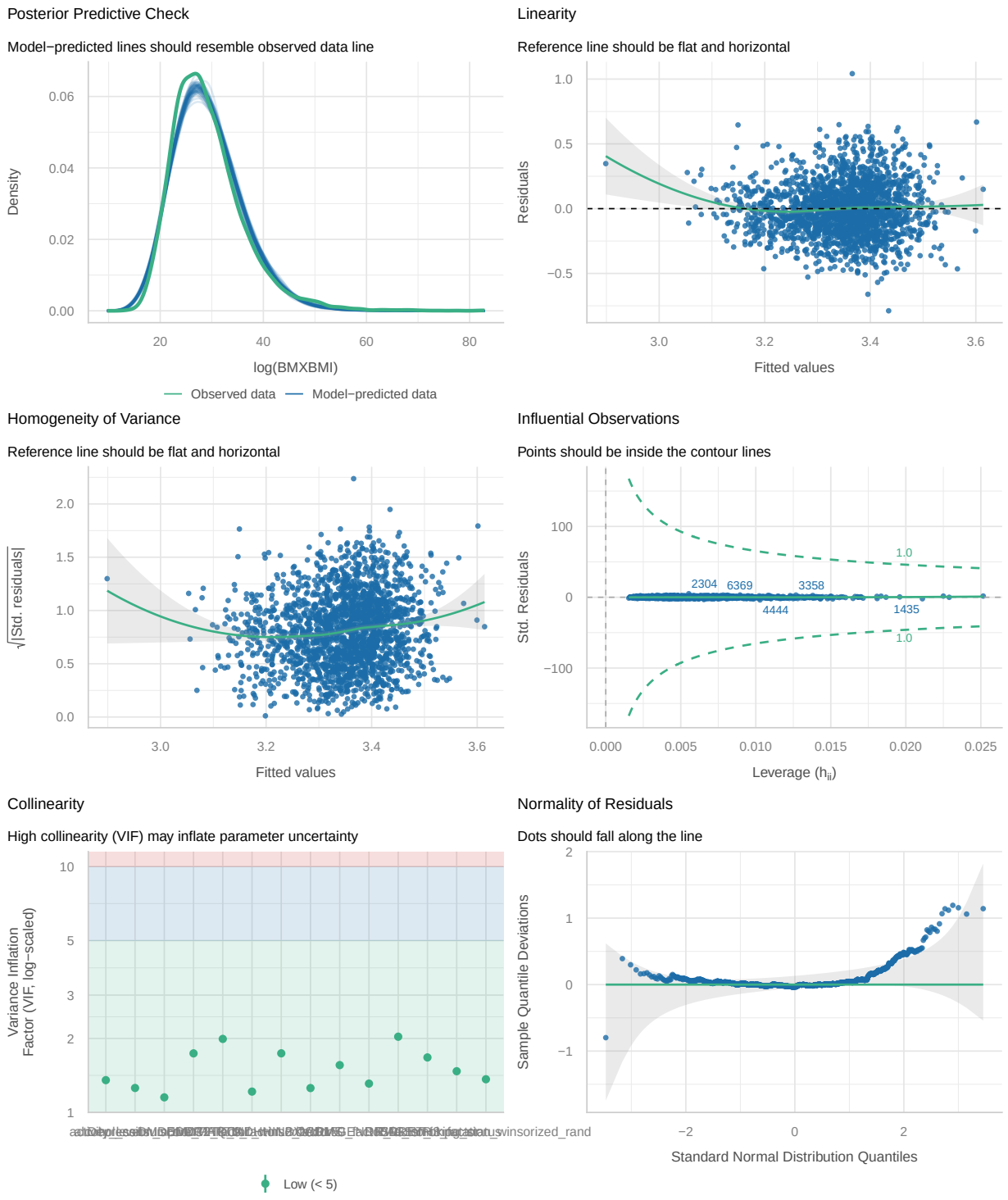

Figure S5: `check_model()`: published SRI.

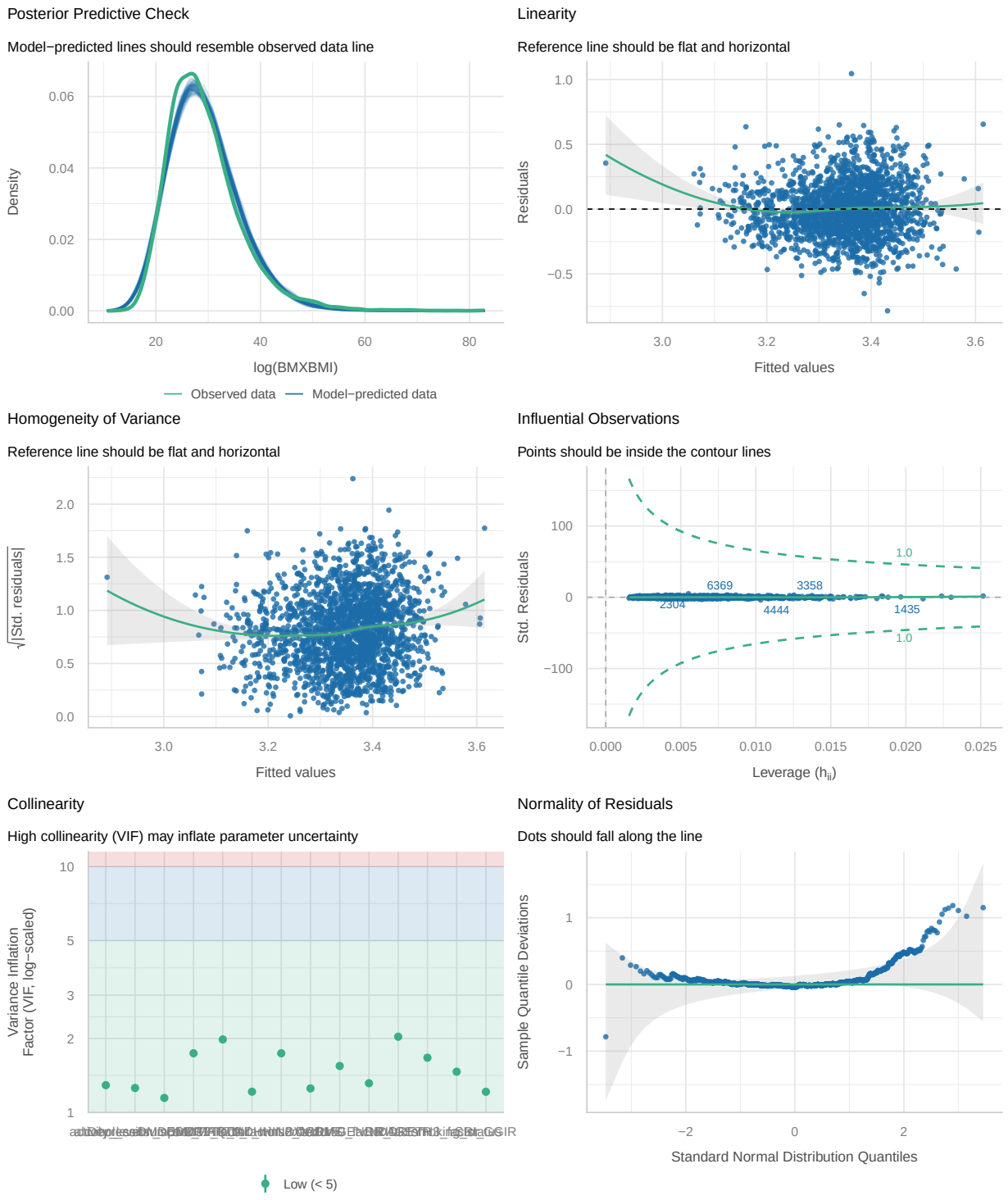

Figure S6: `check_model()`: `GGIR` SRI.

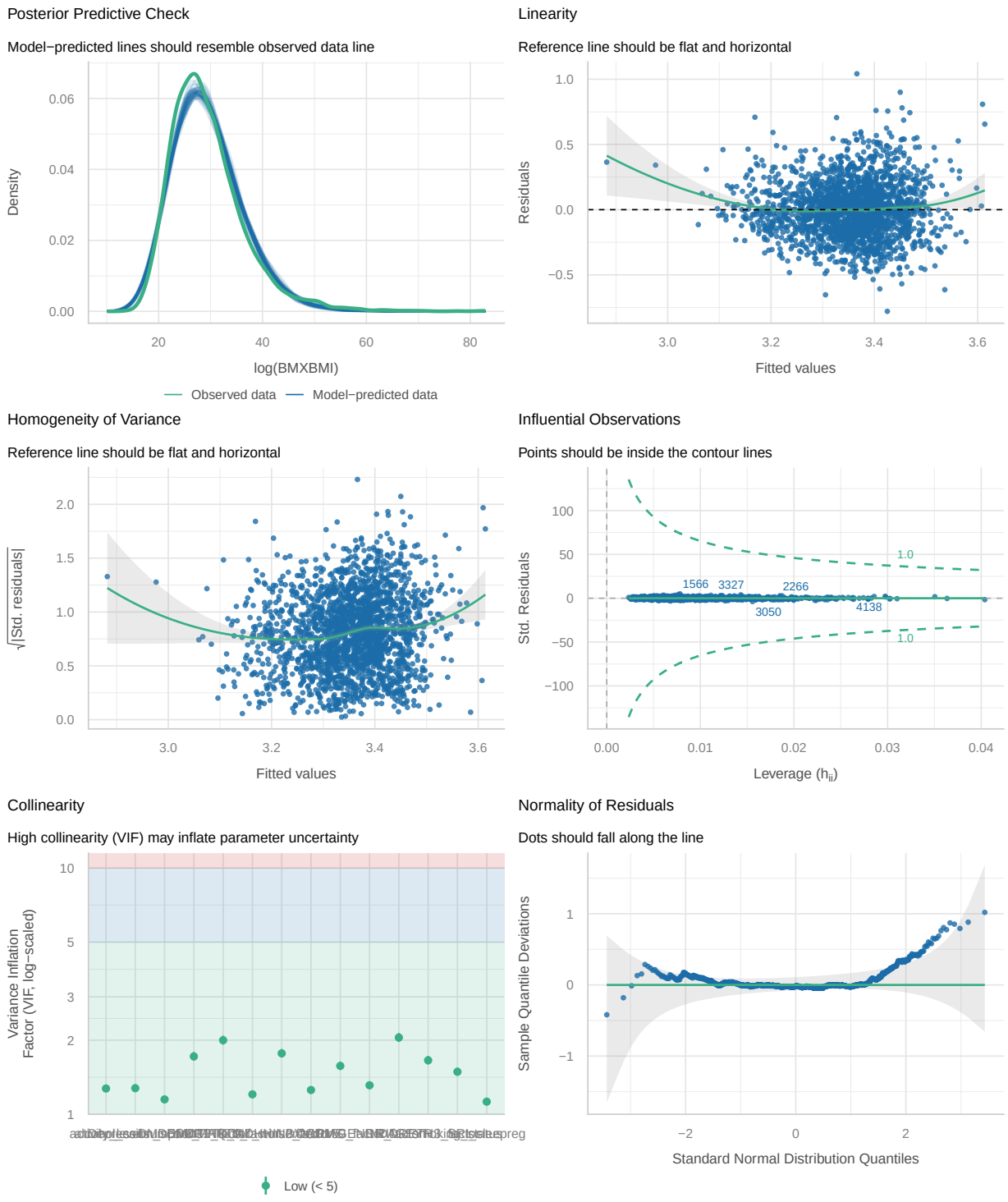

Figure S7: `check_model()`: `sleepreg` SRI.

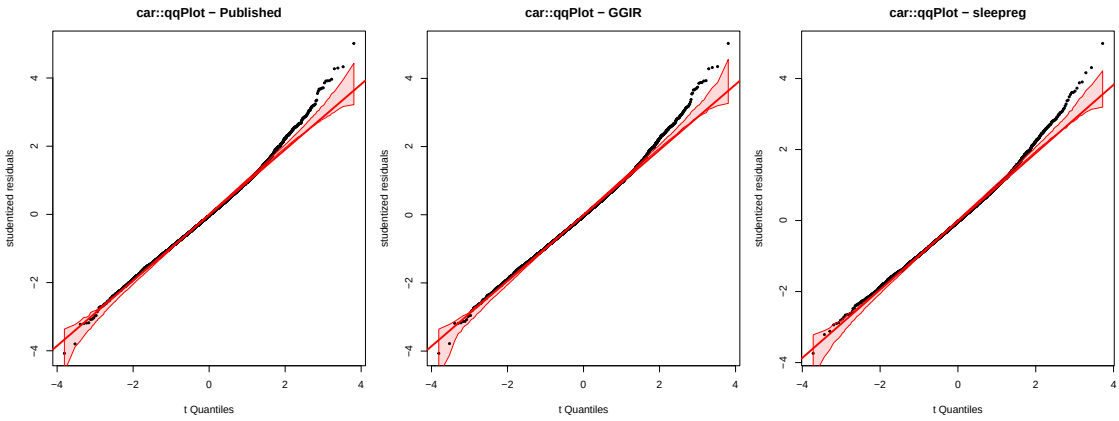

Figure S8: `car::qqPlot()` of studentized residuals with 95% simulated envelopes, fully adjusted model, by calculator. The same mild departure in the upper tail appears for all three.