Robin Wolz[1,2], Janet Munro[1], Ricardo Guerrero[1,2], Derek Hill[1], Yves Dauvilliers[3]
1) IXICO Plc, London, UK, 2) Imperial College London, London, UK, 3) Sleep Unit, Department Neurology, Centre Hospitalier Universitaire, Montpellier, INSERM 1061, France

Extracting digital biomarkers of sleep from 3-axis accelerometry using Deep Learning

Background

- Poor sleep quality is both a risk factor and a symptom of many neurological conditions, including Alzheimer’s disease.
- Wearable biosensors offer practical ways of collecting “digital biomarkers” to identify risk patterns and behaviours in clinical trials and daily life.
- Correct interpretation of the actigraphy data, including clinical context, is essential. Eg. Accounting for age and neurodegenerative disorders.
- We developed and validated a Deep Learning approach to predict sleep from 3-axis accelerometry.

Methodology

- Overnight PSG with simultaneous 3-axis accelerometry was undertaken on 22 elderly subjects from a community study in the Centre Hospitalier Universitaire, Montpellier (age: 85.7±3.7, 15 female, 7 male).
- All subjects wore two Axivity devices (http://axivity.com/) side by side on the same (non-dominant) wrist during the PSG night.
- “Deep Learning” techniques were used to build a convolutional neural network to predict sleep / wake status from accelerometry based on the PSG reference in a leave-one-out approach on the 22 datasets.
- Accuracy for correct sleep / wake classification, as well as an overall sleep efficiency measurement (time asleep / time in bed), were compared between the developed algorithm (referred to as Deep Learning Sleep - DLS) and two standard algorithms employed in current wearable devices: the ESS algorithm [1] that is provided with the Axivity device as well as the widely used Cole-Kripke algorithm. [2].
- A test:re-test analysis was performed for 17 subjects that had usable data from both devices during the PSG night (in the remaining subjects, devices were not deployed correctly or the data was corrupted and not usable).

Results

- Figure 1 shows sensitivity and specificity for detecting sleep with the three algorithms compared.
- Figure 2 shows estimated sleep efficiency as compared to PSG for the three algorithms.
- Test:re-test variability is 3.1%, 6.6% and 5.2% for the Cole-Kripke, ESS and DLS algorithms.
- DLS shows improved accuracy in classifying sleep/awake, compared to both other algorithms. This is most dramatically witnessed in a significant reduction of false positive prediction of sleep and leads to a substantially reduced error between predicted and true sleep efficiency (as measured through PSG).

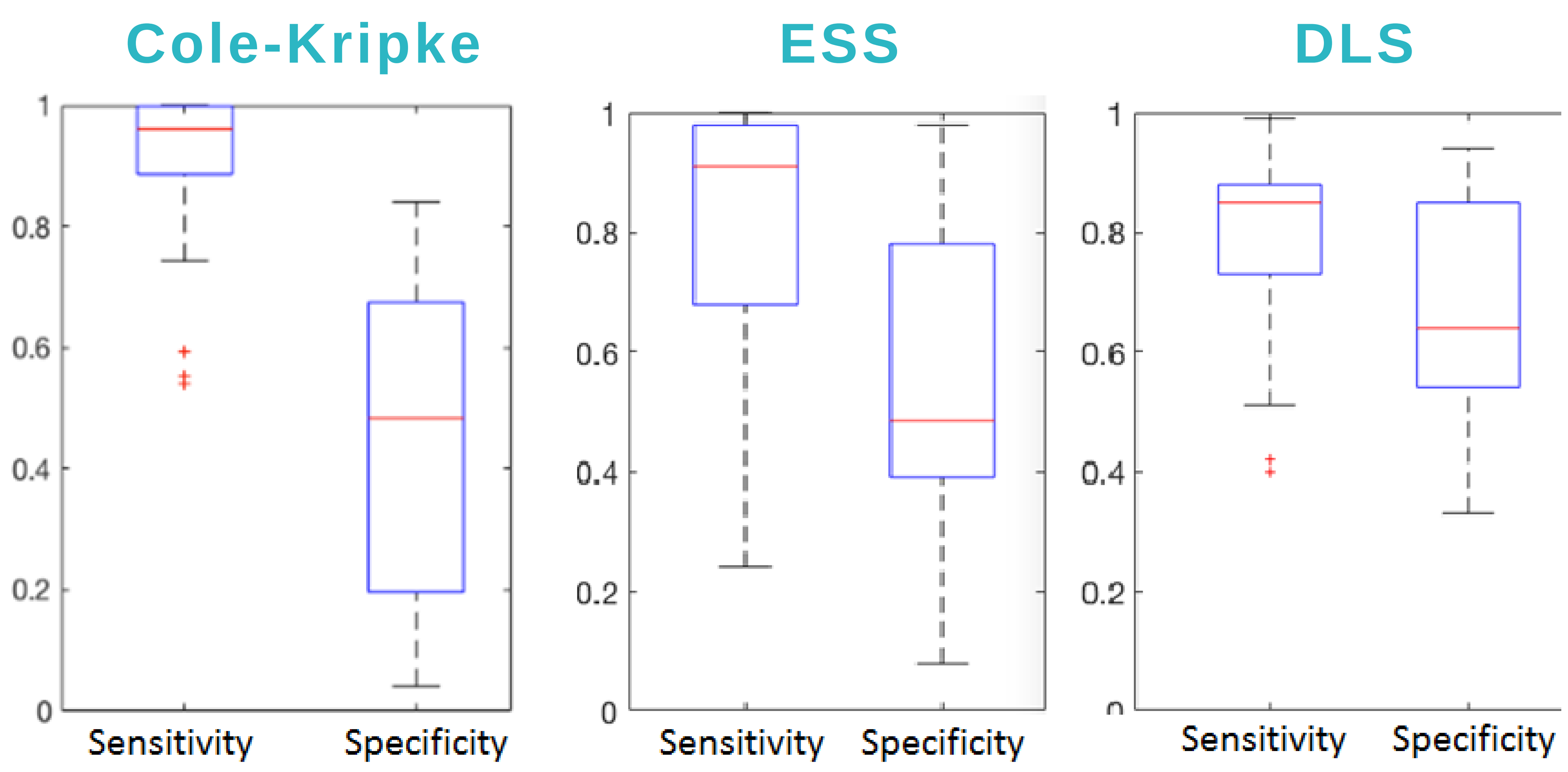

Figure 1: Sensitivity / Specificity to detect sleep

Conclusions

- Deep learning shows potential to identify “digital biomarkers” from wearable sensor data
- The presented algorithm outperforms traditional approaches currently employed in wearable devices
- Periods of quiet (not-moving) wakefulness are common during the night; the standard actigraphy algorithms may be incorrectly interpreting these periods as sleep

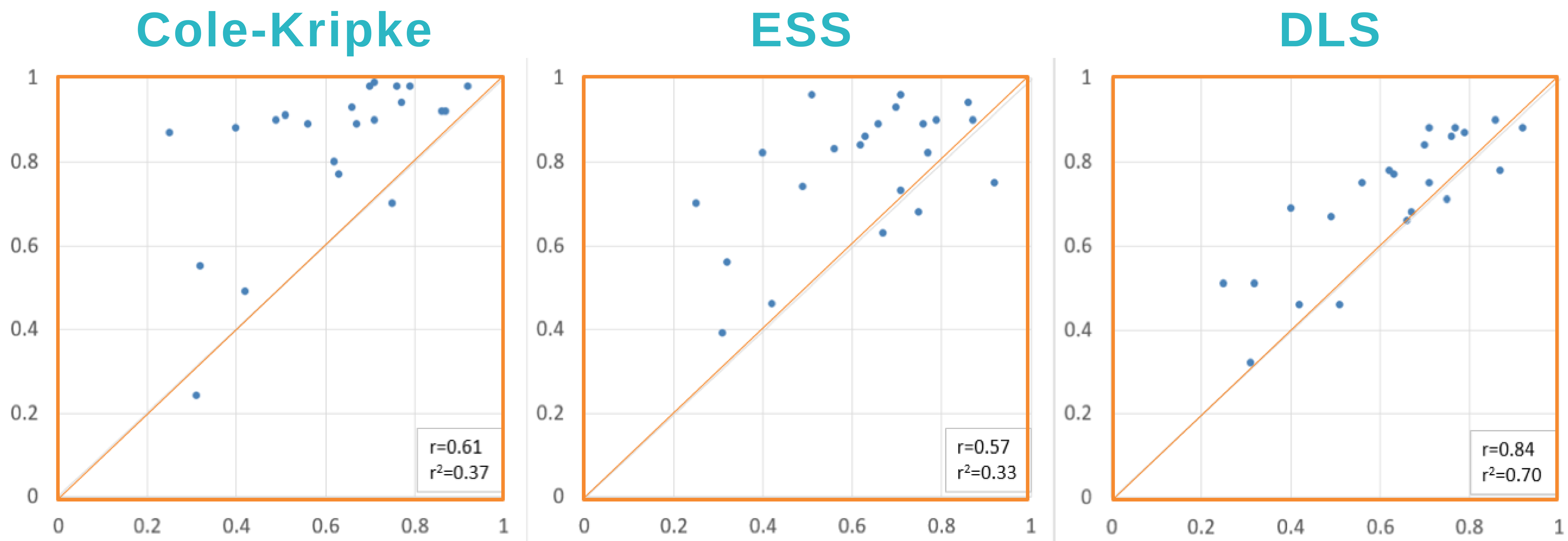

Figure 2: Comparison of sleep efficiency to polysomnography (PSG)

[1] Borazio et al, Towards benchmarked sleep detection with wrist-worn sensing units, International Conference on Health Informatics, 2014:125-134

[2] Cole et al, Automatic sleep/wake identification from wrist activity. Sleep. 1992;15(5):461-9.