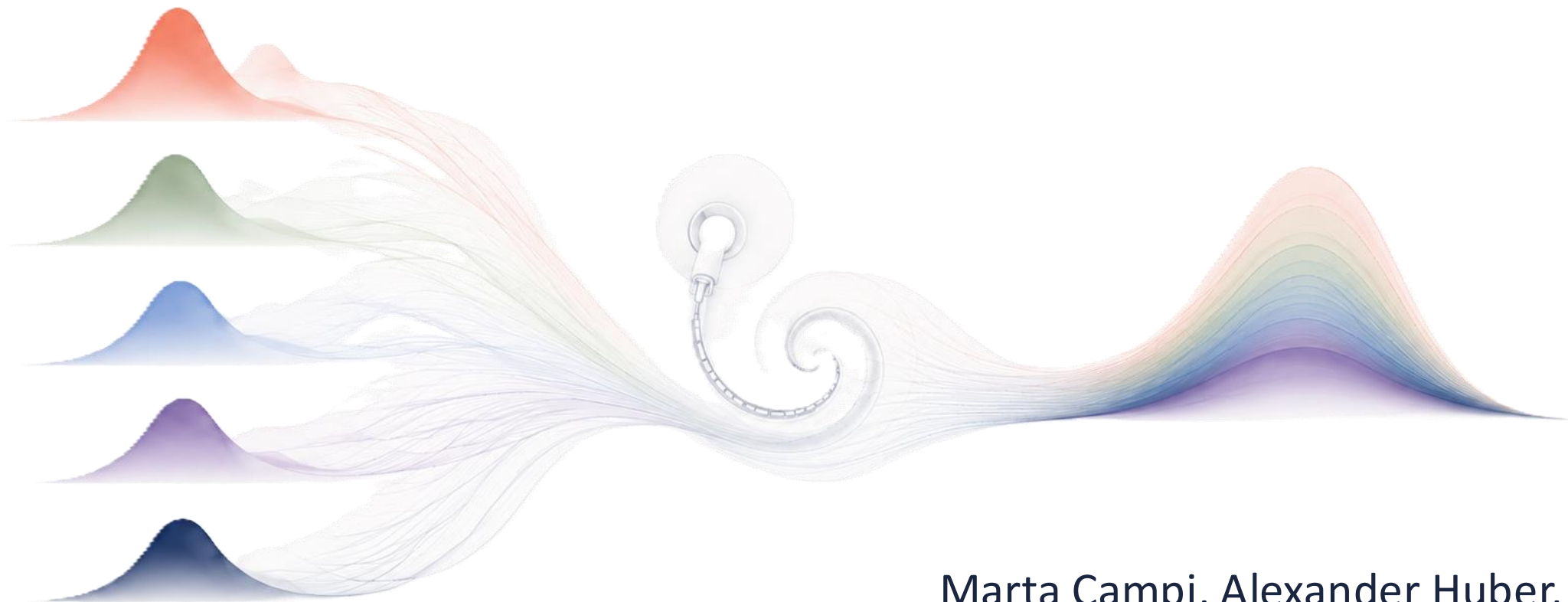


Prediction of Speech Outcomes with Cochlear Implant

Limitations and Opportunities for Clinical Practice



Marta Campi, Alexander Huber, Tobias Goehring

Outcome Prediction in Cochlear Implant (CI)

- *“How well will I hear after the implant?”* is a prediction question
- Needed for **counselling, surgery, monitoring, and rehabilitation.**
- Largest cohorts, every algorithm: error **17-25% points** ^[1,2]; variance explained **< 20%** ^[3]
- So **why** the ceiling and **what** can we predict?



Literature on Outcome Prediction in CI

All studies predict the same outcome: post-operative monosyllabic word score recognition, in quiet, scored as words or phonemes depending on the language

Study	N	Best result (validation)	Most important variables	Model
Kim 2018	120	MAE 4.8→17.1 (LOO* → cross-site)	duration, age, HA use	random forest
Shafieibavani 2021	2,489	MAE 18-22 pp (cross-site)	demographics, audiometry, etiology	ensemble (XGB)
Demyanchuk 2025	2,571	MAE 17 pp (held-out + temporal)	age, duration, pre-op MS, PTA	decision tree
Stronks 2025	533	R ² 0.10-0.19 (in-sample)	duration, age, aided speech	regression
Anthonisen 2026	2,251	R ² = 0.11 (held-out 10×)	age, duration, pre-op hearing	8 models (XGB)
Walia 2022	35	R ² = 0.59 (in-sample)	post-op cochlear measure	regression

MAE = mean absolute error (percentage points); R² = variance explained (0 = no better than the group mean)

* LOO = Leave-One-Out validation

Three Research Questions

1. Prediction of speech outcomes

Regression: can we predict the score?

Dependent variable: Monosyllabic word score recognition (% words correct), continuous

2. Prediction of binary risk for poor outcomes

Classification: can we flag the risk of poor outcome before surgery?

Dependent variable: poor outcome, Monosyllabic word score recognition < 40% (yes / no)

3. Prediction of post-operative trajectories

Forecasting: can we predict it after the first follow-up?

Dependent variable: poor outcome (Monosyllabic word score recognition < 40%) at 12 to 24 months

Cohort: CICH Database, Zurich

- **FULL DATASET**

473 adults, bilateral HL (single-sided deafness excluded) median age 60 (18 to 92) · 53% female

Primary analyses: N = 378 with post-operative Mono1 · follow-up at 0 to 6, 6 to 12, 12 to 24 months

- **PRE-OPERATIVE PREDICTORS** *79 routinely collected variables*

46 Acoustic measures

audiometry · pre-operative speech tests

33 Non-acoustic variables

demographic · developmental · clinical · socioeconomic

- **OUTCOME** *in quiet, 65 dB SPL, free field*

Mono1

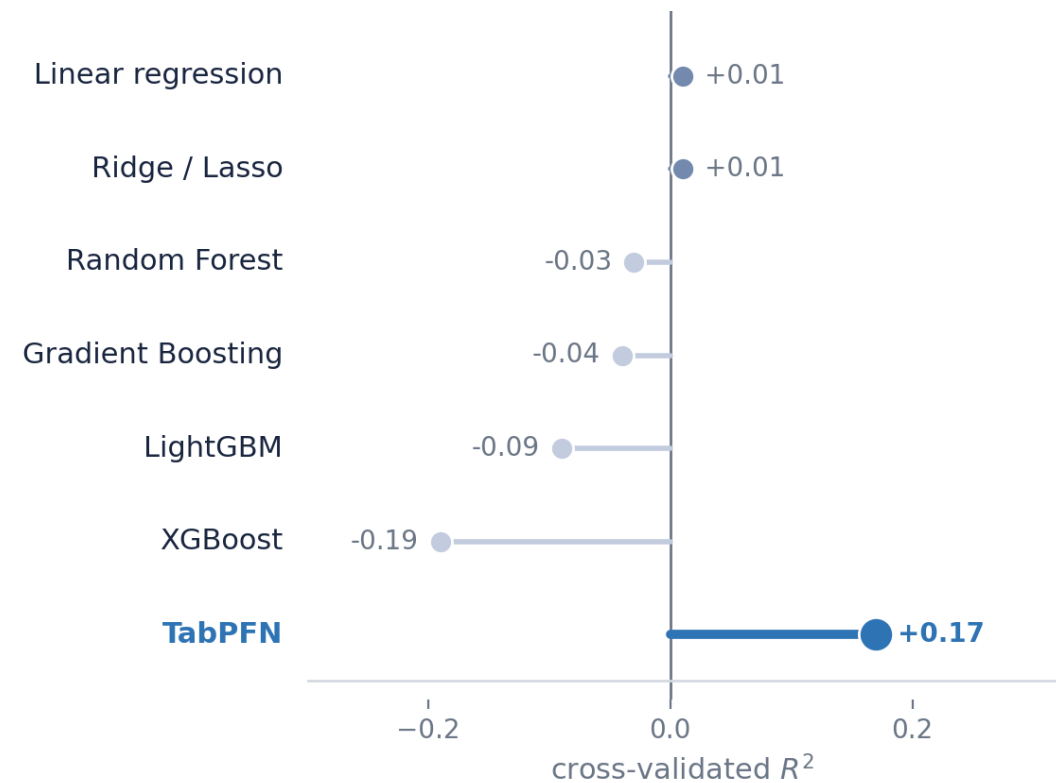
Freiburger monosyllables score measured 12 months after implantation open-set word recognition (% words correct)
Same session: Vowel · Consonant · Voice · Numbers

Poor outcome: **Mono1 < 40%** (about 1 in 5) **measured 12 months after implantation**

Final Score for post-operative trajectories: **Mono1 < 40% after measured after 24 months**

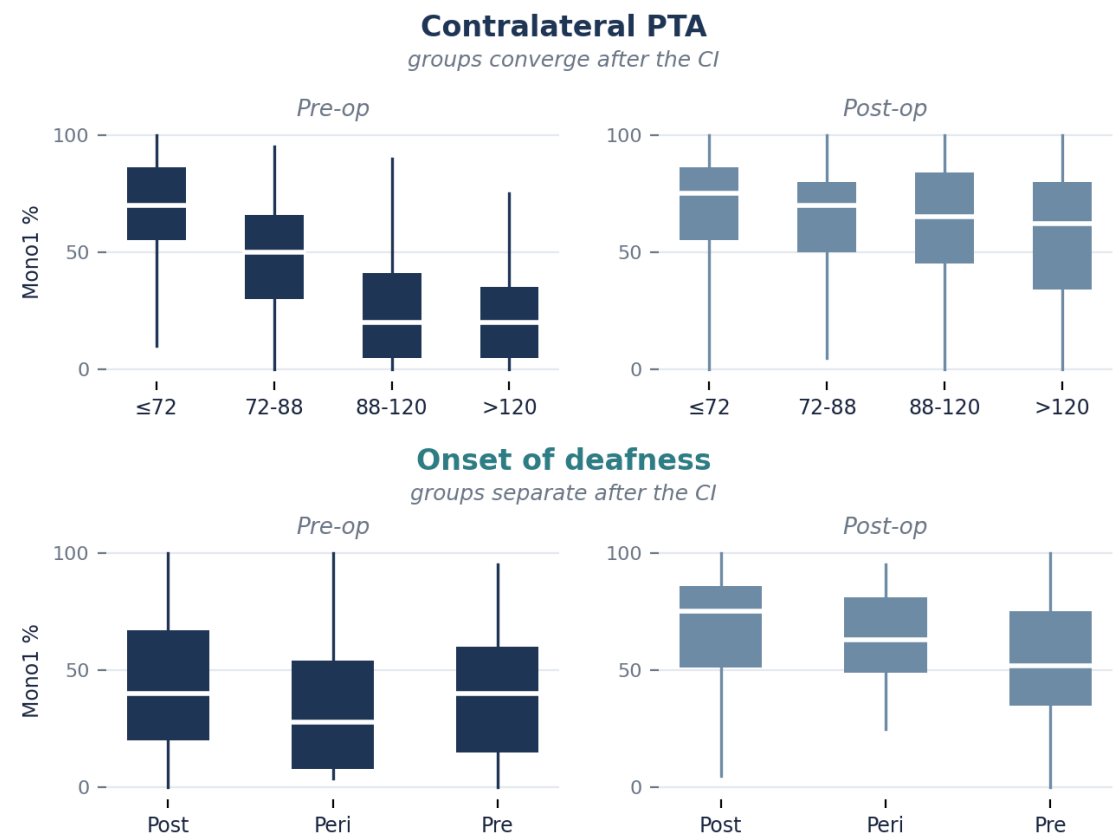
Prediction of Speech Outcomes

EXPLAINED VARIANCE: Best $R^2 = 0.17$



± 41 pp interval (60% means, from 19% to 100%)

WHY? *The CI erases the differences we measured before*

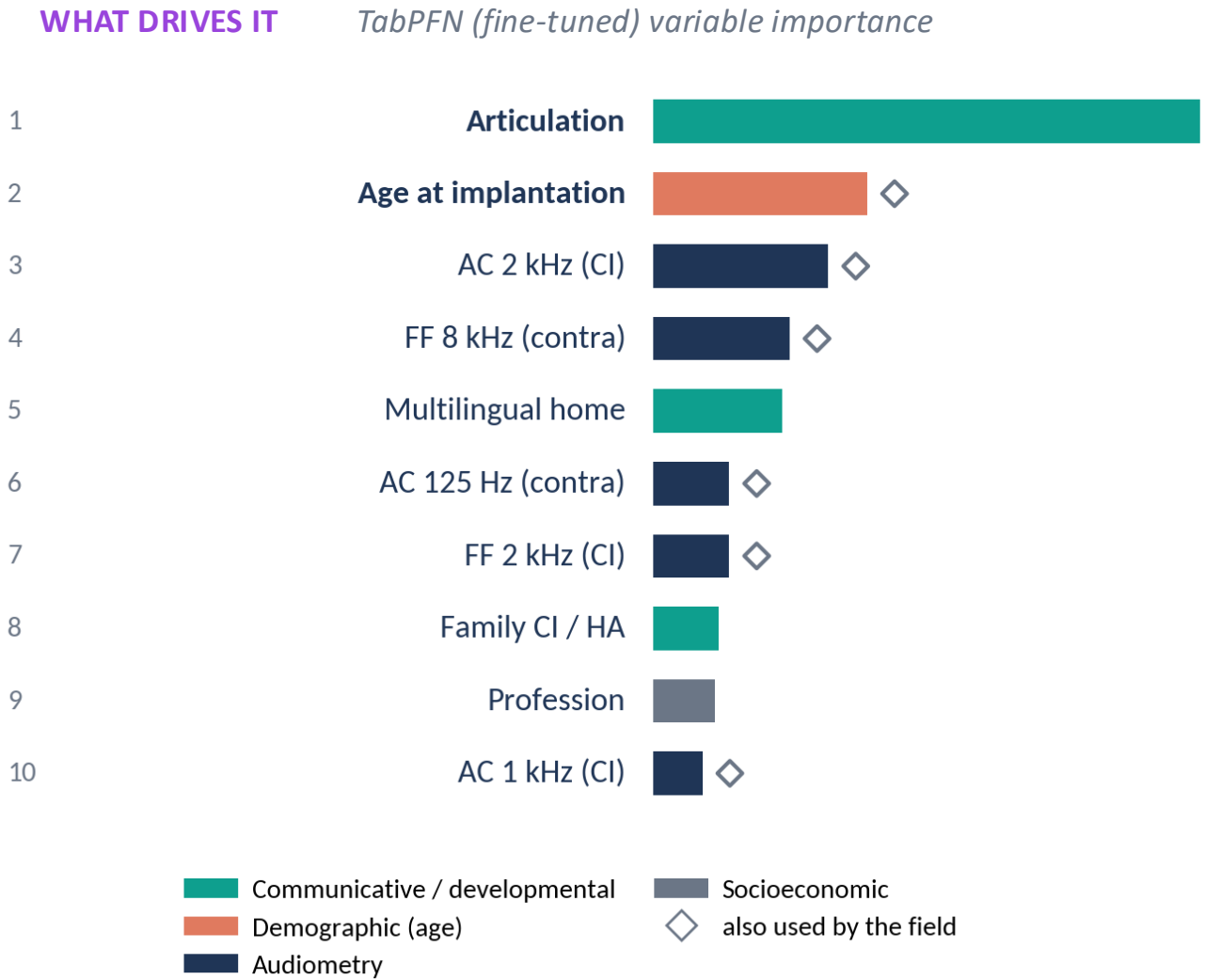


Our pre-op data predict post-op the monosyllabic score with a best $R^2 = 0.17$ with TabPFN (fine-tuned)

Prediction of Binary Risk for Poor Outcomes

- Reframed as **binary risk of poor outcome** (Mono1 < 40%, about 1 in 5 patients)
- TabPFN (fine-tuned), 10-fold cross-validation
- Achieved an Area Under the Curve (**AUC = 0.71**)
- Same binary task, Shew 2025 (N = 1877) reaches AUC 0.71

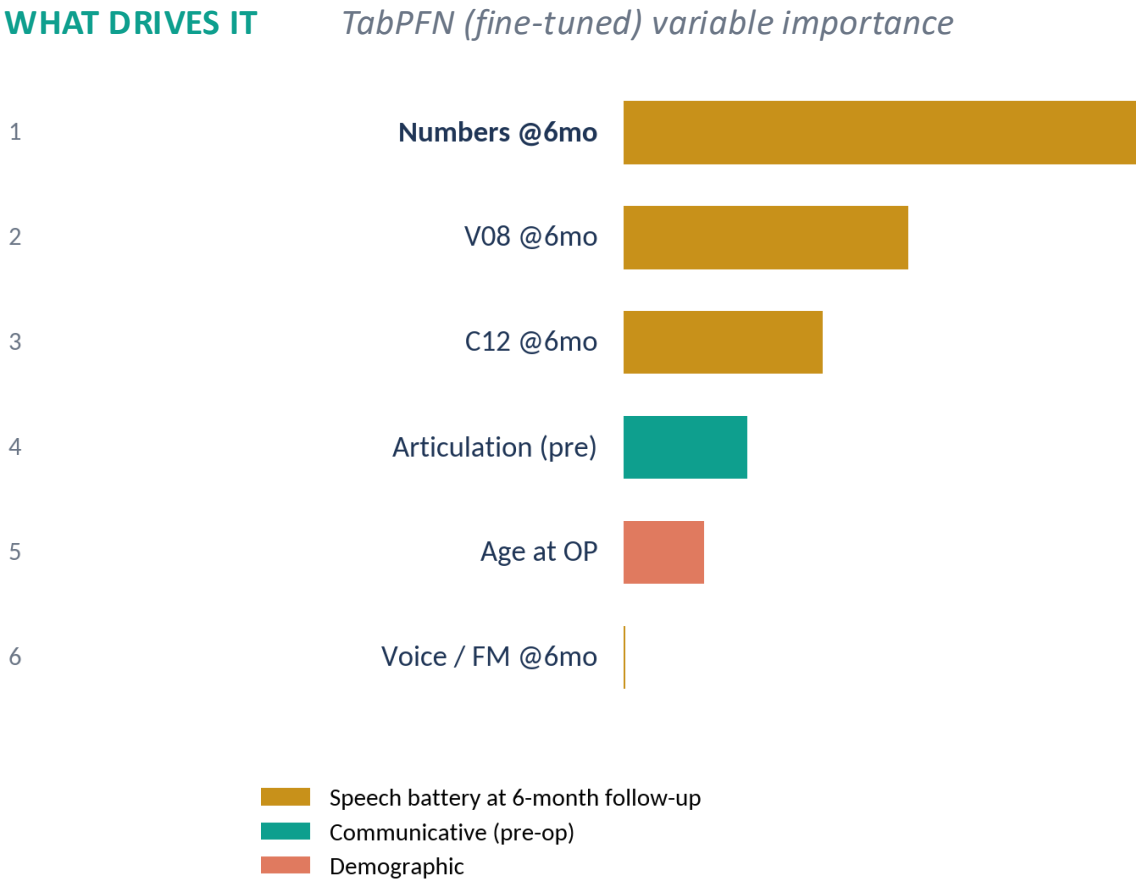
Its strongest driver is **articulation**



◇ = also a top predictor in prior work (Kim 2018, Demyanchuk 2025, Anthonisen 2026, Blamey 2013, Shew 2025). Bars normalized to the top feature; N = 378.

Prediction of Post-Operative Trajectories

- The 6-month follow-up battery forecasts the 12 to 24 month outcome
- It predicts poor Mono1 (< 40%) at **AUC 0.84**, true forward prediction
- Driven by the early battery, mainly number recognition



6-month speech battery (V08, C12, Numbers, voice) predicting poor Mono1 (< 40%) at 12 to 24 months. Permutation Δ AUC, cross-validated, N = 272.

Summary & Conclusions

- Outcomes remain hard to predict, as the field has long known
- In our data, score prediction has a ceiling: $R^2 \leq 0.17$
- As risk it works: **AUC 0.71**, driven by **articulation**
- At the 6-month visit, the speech tests predict the later outcome: **AUC 0.84**
- **Counsel on risk before surgery. Flag poor performers at the first follow-up.**