Forensic experts should focus on measuring and minimising uncertainty, not maximising discriminability

Crises of uncertainty, reproducibility, and replicability in forensic comparison

Vincent Hughes (vincent.hughes@york.ac.uk)
Bruce Xiao Wang (bruce.wang@alumni.york.ac.uk)

Validity: a measure of overall performance, i.e. how well can you separate same- and different-source pairs
Repeatability: intra-examiner/-method/-system reliability
Reproducibility: inter-examiner/-method/-system reliability

Most forensic validation implicitly focuses on

- overall performance of methods under casework conditions – where low values for a given validity metric are the priority
- discriminability (see [5]) with different methods chosen, or decisions made

Forensic comparison is a process with many component parts, involving a variety of subjective and objective decisions, each of which can introduce variability (i.e. uncertainty) into the overall conclusion or performance in a validation

Therefore, the **expert's primary concern should be to recognise those sources of uncertainty, measure that uncertainty, and ideally minimise it** [1,3]

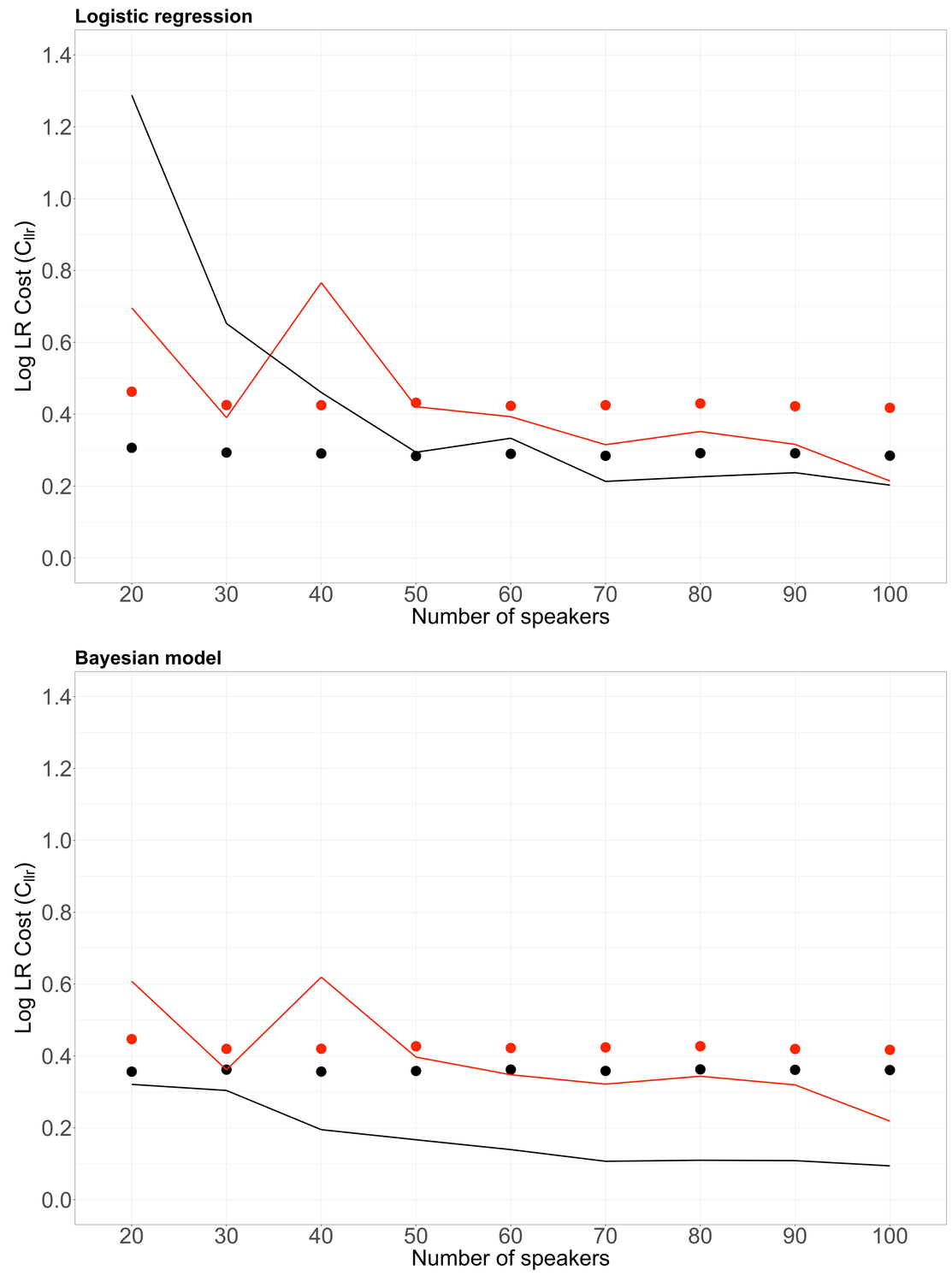
Sources of uncertainty include:

- choice of analytic features, statistical models, reference population, choice of sample represent the reference populations, choice of calibration method
- variation in the unknown (criminal) sample, variation in the known (suspect) sample
- sampling variability
- communication to an end-user (e.g. a judge/jury)

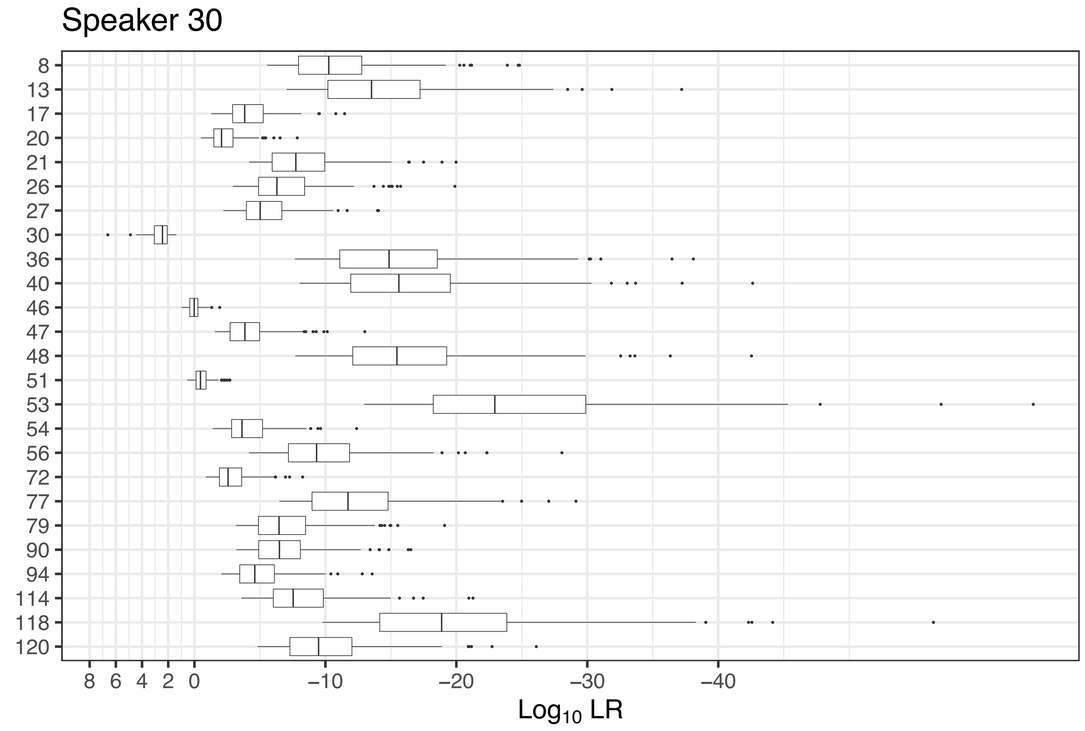
Problem: each case is unique in terms of case materials/conditions

Questions:

- Given how many known and unknown variables there are, how can we estimate uncertainty?
- Is validation even possible given the uniqueness of each case? Are reproducibility and repeatability possible?
- What methods are available for minimising uncertainty?



Speaker 30



Take home messages

- recognise that forensic comparison is a process involving numerous decisions which introduce uncertainty via both systematic and random factors

- be explicit about the decisions made at each stage of the process and the implications of such decisions for uncertainty in terms of the results LRs **and** overall method validity

References

[1] Akmeemana, A., Weis, P., Corzo, R., Ramos, D., Zoon, P., Trejos, T., ... & Almirall, J. (2021). Interpretation of chemical data from glass analysis for forensic purposes. *Journal of Chemometrics*, *35*(1), e3267. [2] Brümmer, N. and Swart, A. (2014) Bayesian calibration for forensic evidence reporting. *Proceedings of Interspeech*, Singapore, 14-18 September. pp. 388-392. [3] Morrison, G. S. and Enzinger, E. (2016) What should a forensic practitioner's likelihood ratio be? *Science and Justice* (Virtual Special Issue on measuring and reporting the precision of forensic likelihood ratios), 56: 374-379. [4] Morrison, G. S. and Poh, N. (2018) Avoiding overstating the strength of forensic evidence: Shrunk likelihood ratios/ Bayes factors. *Science and Justice*, 58: 200-218. [5] Smith, A. M. and Neal, T. M. S. (2021) The distinction between discriminability and reliability in forensic science. *Science and Justice*, 61: 319-331.

UNIVERSITY of York

PP124