Advances in Reliability, Safety and Security

ESREL 2024
Collection of Extended Abstracts

Out-Of-Sample Validation Of Judgemental And Statistical Forecasts Of Antimicrobial Resistance

Abigail R. Colson, Itamar Megiddo

University of Strathclyde, Glasgow, UK

Keywords: expert judgement, expert elicitation, judgemental forecasting, antimicrobial resistance

---

Empirical control is a requirement of advancing science, but it is difficult in the case of expert judgement. How do you validate uncertainty assessments from experts when those assessments are needed because no evidence exists? One approach to expert elicitation, Cooke's Classical Model, overcomes this problem with the introduction of calibration variables (Cooke 1991; Cooke et al., 1988). In addition to asking experts about the variables of interest (i.e., the target questions of the study), the Classical Model also asks experts to assess uncertain quantities from the area of interest that are unknown to the expert, but for which data do exist. Calibration variables allow analysts to validate the expert's performance as an uncertainty assessor by evaluating the hypotheses provided by the expert against the available empirical data. The Classical Model evaluates expert performance on calibration variables in two ways: by measuring the statistical accuracy and the informativeness of the expert's assessments. Analysts can then combine the judgements from multiple experts into a single assessment by weighting the assessments according to expert performance on the calibration variables, with experts who performed well given a high weight and those who performed poorly given a low or zero weight in the overall combined assessment.

The Classical Model has been applied in over 100 real world applications since its inception, and data from previous studies has been compiled into a series of databases collecting studies conducted prior to 2006 (Cooke and Goossens, 2008) and after 2006 (Colson and Cooke, 2017, 2018; Cooke et al. 2021). The availability of this data has allowed researchers to take the notion of validation in the Classical Model one step further. In addition to using performance measures to evaluate individual experts in a given study, the database has been used to evaluate performance-based combinations of experts (as opposed to combining expert assessments by weighting them all equally) more generally, with studies demonstrating both the method's in-sample validity (Cooke and Goossens 2008; Colson and Cooke 2017) and out-of-sample validity, through cross validation methods that split the calibration variables into a training and test set (Eggstaff et al., 2014; Colson and Cooke 2017; Cooke, Marti, and Mazzuchi 2021).

While elicitation studies that enable true out-of-sample validation through the observation of the variables of interest are rare, some such studies do exist. One recent application used the Classical Model to project the proportion of various infections that would be caused by antibiotic resistant bacteria in four European countries (France, Italy, Spain, and the United Kingdom) 2, 5, and 10 years into the future, which corresponded to resistant rates in 2018, 2021, and 2026 (Colson et al., 2019). Data on the observed resistance rates in 2018 and 2021 are now available (European Centre for Disease Prevention and Control 2019; 2022), allowing us to look at the out-of-sample performance of the Classical Model on four panels of experts. Colson et al. (2019) also created a suite of statistical forecasts and found that they did not account for contextual information that experts deemed important when considering future trends in resistance. The updated resistance data for 2019 and 2021 thus also allows us to compare the performance of the expert forecasts with the original statistical forecasts.

References

Colson, A. R., Cooke, R. M. 2017. Cross Validation for the Classical Model of Structured Expert Judgment. Reliability Engineering & System Safety 163 (July), 109–20. https://doi.org/10.1016/j.ress.2017.02.003.

Colson, A. R., Cooke, R. M. 2018. Expert Elicitation: Using the Classical Model to Validate Experts' Judgments. Review of Environmental Economics and Policy 12(1), 113–32. https://doi.org/10.1093/reep/rex022.

Colson, A. R., Megiddo, I., Alvarez-Uria, G., Gandra, S., Bedford, T., Morton, A., Cooke, R. M., Laxminarayan, R. 2019. Quantifying Uncertainty about Future Antimicrobial Resistance: Comparing Structured Expert Judgment and Statistical Forecasting Methods. PLOS ONE 14 (7), e0219190. https://doi.org/10.1371/journal.pone.0219190.

Cooke, R. M. 1991. Experts in Uncertainty: Opinion and Subjective Probability in Science. Environmental Ethics and Science Policy Series. New York: Oxford University Press.

Cooke, R. M., Goossens, L. L. H. J. 2008. TU Delft Expert Judgment Data Base. Reliability Engineering & System Safety, Expert Judgement, 93(5), 657–74. https://doi.org/10.1016/j.ress.2007.03.005.

Cooke, R. M., Marti, D., Mazzuchi, T. 2021. Expert Forecasting with and without Uncertainty Quantification and Weighting: What Do the Data Say?. International Journal of Forecasting 37(1), 378–87. https://doi.org/10.1016/j.ijforecast.2020.06.007.

Cooke, R. M., Mendel, M., Thijs, W. 1988. Calibration and Information in Expert Resolution; a Classical Approach. Automatica 24(1), 87–93. https://doi.org/10.1016/0005-1098(88)90011-8.

Eggstaff, J. W., Mazzuchi, T. A., Sarkani, S. 2014. The Effect of the Number of Seed Variables on the Performance of Cooke's Classical Model. Reliability Engineering & System Safety 121 (January), 72–82. https://doi.org/10.1016/j.ress.2013.07.015.

European Centre for Disease Prevention and Control. 2019. Antimicrobial Resistance Surveillance in Europe 2018. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net). Stockholm: ECDC. https://www.ecdc.europa.eu/sites/default/files/documents/surveillance-antimicrobial-resistance-Europe-2018.pdf.

European Centre for Disease Prevention and Control. 2022. Antimicrobial Resistance in the EU/EEA (EARS-Net): Annual Epidemiological Report for 2021. Stockholm: ECDC. https://www.ecdc.europa.eu/sites/default/files/documents/AER-EARS-Net-2021_2022-final.pdf.