15 June 13:30

Title: AI & ML SIG: Practical AI and Developments in Machine Learning

Sam Hadlington[1], Lesedi Ledwaba-Chapman[2], Paola Berchialla[3], Harry Parr[4], Jason Nicholas[4,5]

[1]Plus-Project Partnership, Hook, United Kingdom. [2]MMS, London, United Kingdom. [3]University of Torino, Turin, Italy. [4]GSK, London, United Kingdom. [5]University of Bath, Bath, United Kingdom

Abstract

AI & ML SIG Updates

A quick discussion on developments within the SIG and our goals for the coming year.

SAP-GPT: Using Modern AI for Document Generation

Large language models (LLMs) could reshape how technical documents are produced in clinical development. This presentation examines the use of LLMs to draft and iterate Statistical Analysis Plans (SAPs) while meeting regulatory and quality requirements, and improving efficiency. We discuss the potential of a practical, end-to-end workflow that grounds generation in the protocol and in a company standard template. Prompt templates and retrieval-augmented generation (RAG) are used to anchor claims to specific protocol sections, preserving traceability.

We will outline the steps we have come up with to generate a SAP and we will talk about the reasoning behind these steps and some issues we ran in to during prompt development. We will show how these steps can be used to build efficiencies into the SAP generation process and how they can be used to make LLM's work for you by having them perform to their strengths. We will also discuss the limitations of LLM's for this particular task that we discovered while building these prompts.

We will then go into a discussion of different LLM's and their relative strengths and weaknesses pertaining specifically to SAP creation.

Finally we will give an assessment of where and when to use this approach and in which situations it could potentially improve development efficiency. We will also discuss how this approach could be applied to other documents and the potential drawbacks of this.

Patient Preference Research with AI-Powered Evidence Generation

Background. Patient Preference Studies (PPS), including Discrete Choice Experiments and other choice-based elicitation methods, are increasingly used to quantify how patients value treatment benefits and risks. However, the synthesis of PPS evidence remains largely manual and fragmented. Heterogeneity in study design and reporting prevents the direct comparison of preferences and limits their integration into regulatory and health technology assessment frameworks.

Objective We developed an AI–based system for the automated extraction, structuring, harmonization, and retrieval of data from published PPS, enabling scalable and reproducible evidence synthesis.

Methods The system integrates OCR and Large Language Models (LLMs), including multimodal architectures, to extract categorical and numerical data from documents, like PDFs, tables, and figures. Categorical variables include study-level descriptors (study type, model specification, disease area), while numerical variables comprise preference weights, p-values, standard errors. Extracted content is validated, harmonized via ontology alignment, and converted into a machine-readable format suitable for meta-analytic aggregation. A real-time module receives natural language queries, interprets intent through vector search and LLM-based semantic parsing, and retrieves the corresponding harmonized data.

Results The system replaces manual data extraction and expert-driven normalization, achieving consistent, reproducible structuring of PPS outputs. Our system was validated using the SIGMA

framework, which classifies outputs as correct, incomplete, wrong, missing, or hallucinating. Overall, 89% of extracted data were correct, 6.5% incomplete, 4.2% wrong, no missing or hallucinating outputs.

Conclusions This framework represents the first end-to-end AI solution for automated extraction and on-demand retrieval of PPS evidence, enabling scalable, transparent, and regulatory-ready meta-analysis of patient preferences.

A Tutorial on Tidymodels

The development of prediction models, e.g. statistical and machine learners (ML), are fast becoming standard practice within life sciences and medicine. These learners aim to predict outcomes / prognosis which can guide clinical decision making, e.g. support further treatment interventions. Widely used examples in clinical practice include the Framingham risk score for estimating 10-year cardiovascular risk . Predictive models are also being integrated into prospective clinical trials, to improve precision in estimating the treatment effect by complementing standard covariate adjustment. For instance, ML-derived prognostic scores can better capture nonlinear relationships, supplementing trial efficiency.

At the same time, the life sciences sector is rapidly adopting more open-source software, leveraging ecosystems like the tidyverse as well as more bespoke packages under the pharmaverse, which promotes best practices within the industry. This momentum in open-source adoption and the utility of robustly developed prediction models brings recent work in tidymodels into sharp focus. We will present a concise primer, introducing a multi-step unified framework for developing and validating prediction models using the tidymodels R package.

Tidymodels offers a comprehensive ecosystem for building predictive models in R. The workflow spans data preprocessing via ‘recipes’, models specification (including ensembles), and resampling & hyperparameter tuning, where we combine these steps into a unified pipeline for evaluating the performance metrics (e.g. R^2 and other loss-functions). Visual and explainable methods will also be demonstrated. The session will be useful for attendees who wish to see practical implementations of complete modelling frameworks within R.

Panel Discussion

All of the speakers will come together at the end on a panel to field any questions from the audience.