AI-Assisted Generation of Synthetic Patient-Level Data to Recreate Weight-Change Outcomes from the STEP-4 Semaglutide Trial

Massachusetts College of Pharmacy and Health Sciences

Devendra Patil, MS, PhD Candidate; Hannah Vuong, PharmD Candidate; Joanne Doucette, MS, MLS; and Tewodros Eguale, MD, MS, PhD

MSR50

BACKGROUND

- The STEP-4 trial is a 68-week trial that enrolled adults with obesity or overweight to evaluate the effect of continuing versus withdrawing semaglutide treatment.
- Because patient-level data (PLD) from randomized controlled trials are not publicly available, opportunities for secondary analyses are limited.
- Large language models (LLMs) can help address this gap by generating synthetic PLD from published summary data, enabling reproducibility, visualization, and exploratory investigation.

OBJECTIVE

- To evaluate whether an AI model (ChatGPT) can generate realistic synthetic PLD and, after simulating dropout patterns, accurately reproduce the weekly weight-change trajectory and participant counts reported in the STEP-4 trial

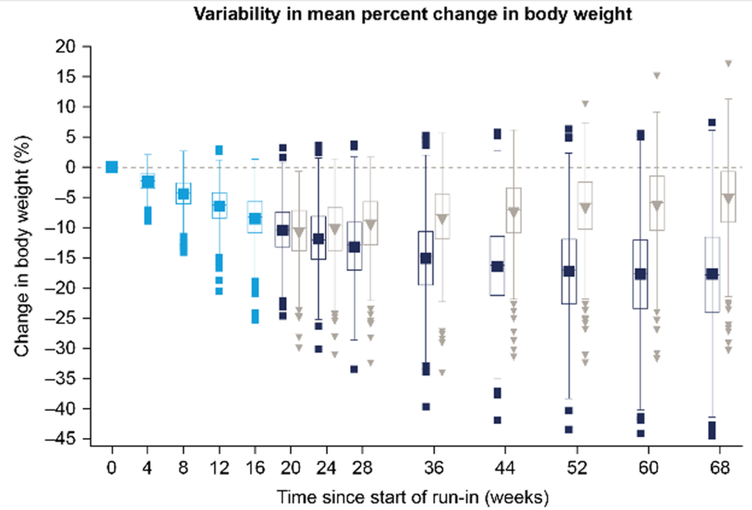
METHODS

- Data was extracted from published STEP-4 Trial (Semaglutide vs placebo)
- Synthetic data was generated using ChatGPT (summary stat and graphics)
- Participant dropout was simulated to match weekly sample sizes
- Change in body weight and 95% CI was calculated using SAS software

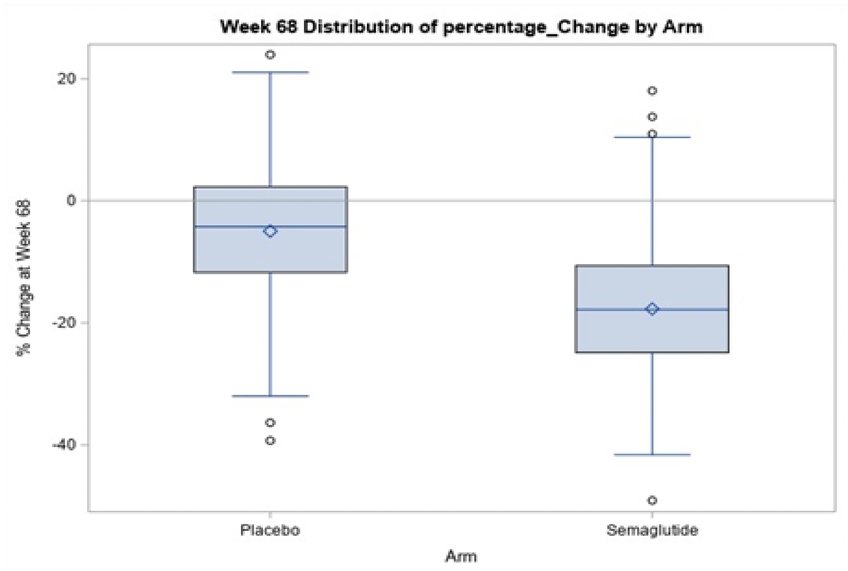
TABLE 1: Mean percentage weight change at Week 68 with corresponding 95% confidence intervals by treatment arm (synthetic patient-level data)

Arm	Week	Mean (%)	95% CI (Lower)	95% CI (Upper)
Semaglutide	20	-10.03	-10.62	-9.44
Placebo	20	-10.96	-11.84	-10.08
Semaglutide	36	-12.59	-13.20	-11.98
Placebo	36	-8.97	-9.91	-8.03
Semaglutide	52	-15.15	-15.85	-14.44
Placebo	52	-6.98	-8.05	-5.91
Semaglutide	68	-17.70	-18.55	-16.85
Placebo	68	-4.99	-6.25	-3.73

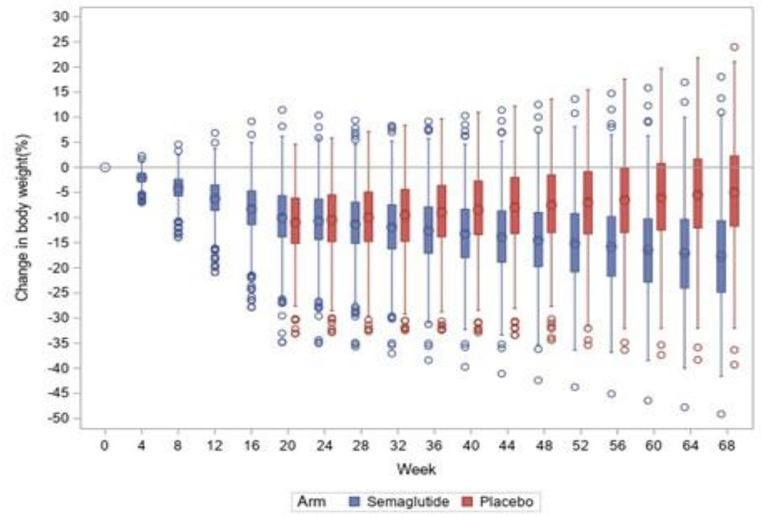
GRAPH 1: Mean percentage weight change at Week 68 with 95% confidence intervals by treatment arm reported in STEP-4 Semaglutide Trial



GRAPH 2: Distribution of percentage weight change at Week 68 by treatment arm based on synthetic patient-level data



GRAPH 3: Mean percentage weight change at Week 68 with 95% confidence intervals by treatment arm plotted- synthetic patient-level data



RESULTS

- Week 68 estimates mean percentage weight change closely matched: Published results (-17.4% Semaglutide vs -5.0% Placebo) & Synthetic data (-17.7% Semaglutide vs -4.99% Placebo)
- This approach demonstrates the feasibility of conducting exploratory analyses, including subgroup analyses based on baseline weight, using synthetic data when original patient-level data are unavailable

KEY FINDINGS

LLM-generated synthetic data can reproduce RCT outcomes and uncertainty without access to patient-level data.

LIMITATION

The absence of detailed patient-level BMI data and longitudinal BMI reporting in the original publication limited the ability to generate reliable synthetic BMI trajectories. This constraint restricted subgroup analyses based on BMI and limited deeper evaluation of heterogeneity of treatment effects.