Supplementary Text S1. Standardized Patient-Specific Prompt Used for GPT-5.6 Sol Predictions

The following is an English translation of the standardized Turkish-language prompt used to generate patient-specific lumbar spine and total hip bone density score estimates. The example contains the demographic, anthropometric, clinical, and biochemical values of an illustrative patient.

Standardized prompt (English translation):

"Using the demographic, anthropometric, clinical, and laboratory data of a 45-year-old premenopausal woman with a height of 160 cm, weight of 78 kg, body mass index (BMI) of 30.4 kg/m², serum 25-hydroxyvitamin D level of 20.3 ng/mL (vitamin D insufficiency), calcium level of 10.1 mg/dL, phosphorus level of 3.3 mg/dL, parathyroid hormone (PTH) level of 65 pg/mL, creatinine level of 0.7 mg/dL, alkaline phosphatase (ALP) level of 47 U/L, and no history of previous fractures, estimate the quantitative computed tomography (QCT)-derived bone mineral density scores for the lumbar spine (L1–L4) and total hip.

Use T-scores for postmenopausal women and men aged 50 years or older, and Z-scores for premenopausal women and men younger than 50 years.

Generate one numerical estimate for each anatomical region based exclusively on the provided patient information.

Report the results only in the following format, without any additional explanation:

Lumbar spine (L1–L4) QCT Z/T-score: …

Total hip QCT Z/T-score: …"

Prompt administration: The same standardized prompt structure was used for all 450 patients, with patient-specific demographic, anthropometric, clinical, and laboratory values substituted for each assessment. Each patient was evaluated through a single query in a separate ChatGPT conversation. No preliminary prompt testing or iterative prompt refinement was performed before or during the evaluation of the study cohort.