Algorithm 1 Three-Stage OMOP CDM Entity Mapping

procedure EntityMapping($entity, domains$)
  $E \leftarrow \text{SapBERT}(entity)$
  for $d \in domains$ **do**
    $C_1 \leftarrow \text{Stage1}(entity, d, E)$
    $C_2 \leftarrow \text{Stage2}(C_1)$
    $C_3 \leftarrow \text{Stage3}(entity, C_2, E)$
    $results[d] \leftarrow C_3[0]$
  end for
  return $\arg\max_d results[d].score$
end procedure

procedure Stage1($entity, domain, E$)
  $C \leftarrow \text{ES.textSearch}(entity, domain, 5)$ ▷ Lexical
  $C \leftarrow C \cup \text{ES.knnSearch}(E, domain, 5)$ ▷ Semantic
  $C \leftarrow C \cup \text{ES.hybridSearch}(entity, E, domain, 5)$ ▷ Combined
  return C ▷ 15 candidates
end procedure

procedure Stage2(C_1)
  $S \leftarrow \{c \in C_1 \mid c.std \in \{\text{'S'}, \text{'C'}\}\}$
  for $c \in C_1$ where $c.std \notin \{\text{'S'}, \text{'C'}\}$ **do**
    $rels \leftarrow \text{ES.search}(\{id_1 : c.id, rel : \text{"Maps to"}\})$
    $S \leftarrow S \cup \{r.target \mid r \in rels, r.target.std \in \{\text{'S'}, \text{'C'}\}\}$
  end for
  return Deduplicate(S)
end procedure

procedure Stage3($entity, C_2, E$)
  for $c \in C_2$ **do**
    $s_{text} \leftarrow \begin{cases} 1.0 & \text{if } c.is\_mapped \\ \frac{|N_3(entity) \cap N_3(c.name)|}{|N_3(entity) \cup N_3(c.name)|} & \text{otherwise} \end{cases}$
    $s_{sem} \leftarrow \frac{(E \cdot c.emb)/(\|E\| \|c.emb\|)+1}{2}$
    $c.score \leftarrow 0.4 \cdot s_{text} + 0.6 \cdot s_{sem}$
  end for
  return SortDesc(C_2)
end procedure