



# NCCD – RxNorm: Linking Chinese Clinical Drugs to International Drug Vocabulary

Yaoyun Zhang, PhD<sup>1</sup>, Jing Li<sup>2</sup>, Mui Van Zandt<sup>2</sup>

1 Melax Technologies, Inc, 7000 Fannin St Suite 1900, Houston, TX, USA; 2 Real World Solutions, IQVIA, Durham, NC, USA

## Background

The emerging pandemics make it urgent to establish a clinical research network on a global scale. The expansion of clinical research from China to the international community will greatly facilitate cooperative efforts. However, due to the language barrier and different vocabulary systems, important clinical knowledge in China remains unexchangeable with the other countries. This project aims to map concepts of chemical drugs in a Normalized Chinese clinical drug (NCCD – Figure 1) knowledge base with RxNorm and the international RxNorm extension in OHDSI research network, and build an efficient channel for clinical research both locally and globally.

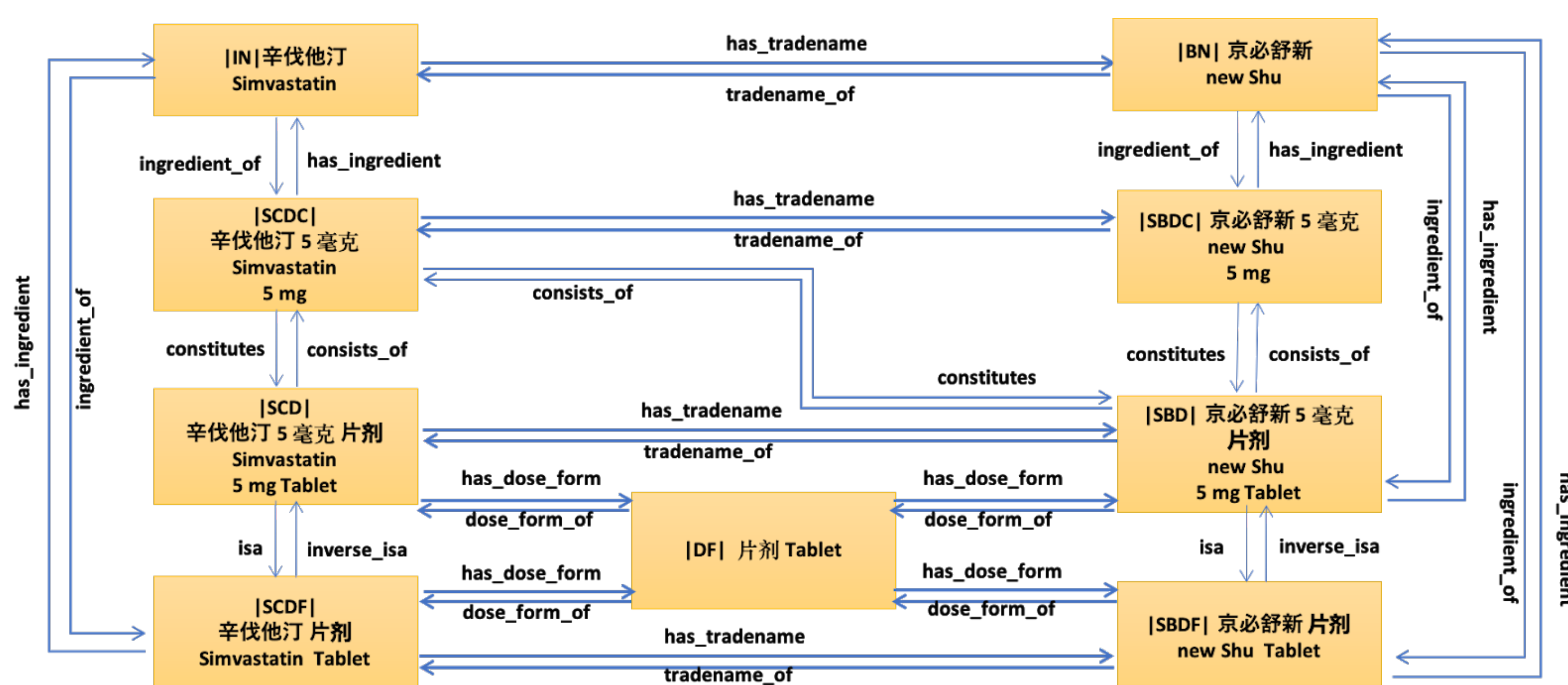


Figure 1. Information model of chemical drugs in NCCD

## Methods

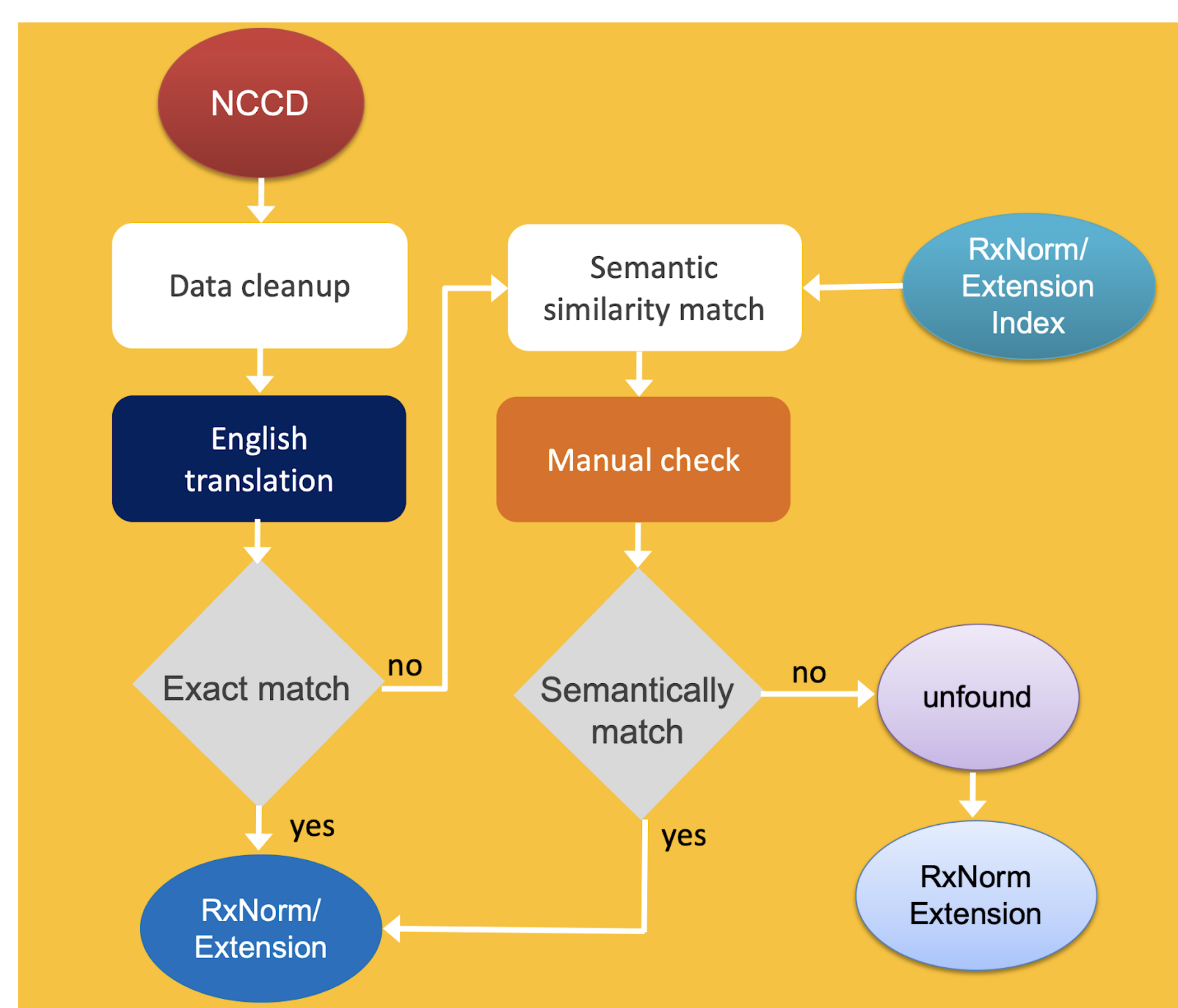


Figure 2. Workflow of semantic mapping between NCCD and RxNorm/Extension.

The workflow is illustrated in Figure 2:

- Data cleanup:** remove concepts of Chinese Patent Drugs and noises of mislabels;
- Concept translation:** Terms translated into English, by using known Chinese-English term pairs, manual check based on auto-translators;
- Concept matching:** To process English translations that cannot find exact string match in RxNorm terms, a search engine with indexed RxNorm terms was built. The most semantically similar RxNorm terms will be returned for manual review;
- Enrich RxNorm-Extension:** the rest concepts in NCCD that cannot find any match will be added into RxNorm-Extension.

## Results

The statistics of current mapping results is illustrated in Table 1. In total, 2 159 ingredients (64.83%), 931 dose forms (90.92%) and 88 branded names (1.80%) in NCCD can be mapped to RxNorm. 2,247 (25.39%) and 588 (6.64%) semantic clinical drugs in NCCD can be mapped to RxNorm and RxNorm Extension, respectively. We further validated the mapping results using top 1,000 Chinese drugs frequently used in clinical settings. Specifically, as illustrated in Figure 3, among the top 1,000 most frequent clinical drugs in China, 673 (67.3%) drugs are mapped to RxNorm at the ingredient level, 358 and 52 are mapped to semantic clinical drugs in RxNorm and RxNorm Extension (41.0%).

Table 1. Statistics of semantic mapping between NCCD and RxNorm/Extension.

Chemical Drug Concept in NCCD	Mapping to RxNorm/Extension
Ingredient (IN)	2,159 / 64.83%
Dose form (DF)	931 / 90.92%
Brand name (BN)	88 / 1.80%
Semantic clinical drug form	3,075 / 37.63%
Semantic clinical drug component	788 / 9.64% (Extension)
Semantic clinical drug	2,226 / 33.06%
	578 / 8.58% (Extension)
Semantic clinical drug	2,370 / 22.70%
	588 / 5.63% (Extension)

Table 2. Examples of mapping frequent clinical drugs in prescriptions to NCCD and RxNorm/Extension.

Normalized clinical drugs	Translate	RxNorm Concept	RxNorm CUI	Semantic type
盐酸曲美他嗪 20 mg 片	trimetazidine 20mg tablet	trimetazidine	10826	IN
阿司匹林 25 mg 肠溶片	aspirin 25 mg enteric-coated tablet	aspirin 25 mg	329295	SCDC
华法林 3 mg 片剂	warfarin 3mg oral tablet	warfarin oral tablet	374319	SCDF
阿托伐他汀 20 mg 片剂	atorvastatin 20 mg oral tablet	atorvastatin 20 mg oral tablet	617310	SCD
缬沙坦 80 mg 胶囊剂	valsartan 80 mg oral capsule	valsartan 80 mg oral capsule	199850	SCD

## Conclusions

The mapping between NCCD to RxNorm/Extension will serve as an efficient channel to promote clinical research both locally and globally. Studies using real-world data will be conducted in the near future to further validate and improve the usability of the generated drug information resources.