SAS
GLOBAL
FORUM
2019

USERS PROGRAM

APRIL 28 - MAY 1, 2019 | DALLAS, TX
Introduction

At least three human vaccine quality scandals broke out in China in the past 5 years. The recent one that was reported in July 2018 triggered an outcry from Chinese parents on the internet. They were eager to find out whether the vaccines that their children got were at risk. However, there was no complete vaccine tracing database that was available to public for easy search. We gathered official vaccine sourcing information and consolidated the data. We used Base SAS for web scraping, data cleansing, and data consolidation. Our database is used to provide an easy way to check the risk of human vaccines that were in the market in the past years.

Data Source

Our data source is public government procurement information. In China, all government procurement information should be made public on the provincial level. We gathered information for each province separately. Most provinces have more than one release about vaccine procurement available on the internet. The provincial data was then consolidated into a national-level multi-year database.

Application

The final database is used for website wodeyimiao.net. On this website, users can check their vaccine source and respective quality risk level by province name, vaccine name, or/and vaccination year.

Our data collection will be an on-going process and we will keep gathering future vaccine procurement data to include into this database.
Tracking and Consolidating Chinese Vaccine Source Information

Yizhe Ge, Tracy Song-Brink
Donghua University, North Carolina State University

Example:

SAS code to get information from HTML:
filename sourcing url
data hubei2017;
  infile sourcing length=len lrecl=32767;
  input line $varying32767. len;
  output;
run;
SAS GLOBAL FORUM 2019

APRIL 28 - MAY 1, 2019 | DALLAS, TX

Kay Bailey Hutchison Convention Center