

# Development and Application of the CFSAN Analytical Product Survey (CAPS) Database: FDA's Tool for Effectively Managing Marketplace Survey Data



Cynthia Srigley, Edward Appiah, James Sanders, Marla Surette, Arsh Randhawa, Sheela Gopalakrishnan, Jennifer Paine, Thomas Farrell, Gregory Noonan, Christine Parker  
Center for Food Safety and Applied Nutrition, Food and Drug Administration, College Park, MD

## Abstract

CFSAN's Office of Regulatory Science/Division of Bioanalytical Chemistry (ORS/DBC) routinely performs sampling studies, or surveys, of foods and dietary supplements from the United States (US) market to collect information on ingredients and contaminants. ORS/DBC, in collaboration with CFSAN's Office of Management/Division of Information Technology (OM/DITM), developed an internal, relational database to store, analyze, and communicate information related to sampling studies performed or coordinated by CFSAN. The CFSAN Analytical Product Survey (CAPS) database provides a standardized platform for managing product label information (e.g., ingredients, nutrients, and serving size information) and analytical data generated for each sampled product. Data are available to users through an interactive web-based interface with data visualizations (Dashboard), product information tables (Data View), and computational calculations (Reports). Filtering, quick search, and referral level flagging capabilities enhance the customized data analysis experience. As a benefit to the Agency, the CAPS database improves communication with stakeholders, provides a platform for standardized reporting, promotes data integration and interoperability, helps to inform signal detection, and supports feasibility assessments. The CAPS database currently houses information for two cannabis-derived product sampling studies, including data on 1,519 products and 40,000 data points, with expansion to new analytical product surveys already underway. The CAPS database provides a user-friendly, interactive tool for effectively managing big data and builds capacity for future data uses within CFSAN.

## Introduction

### Statement of Business Need

In the absence of a unified data system, sample data generated by CFSAN (ORS or third-party laboratory) have been historically managed by individuals, with data collection and analysis tailored to specific use cases.

As data sources and volumes grow, data governance strategies are required to ensure that data assets are from quality sources and managed consistently throughout the enterprise. Directed by the mission to modernize data sources and expand data capacity through integration and interoperability, the CAPS database was first conceptualized using product data from two cannabis-derived product sampling studies, with ambitions to encompass any product survey performed by CFSAN/ORS in support of program office data needs.

### New Survey Checklist

The following checklist is used to determine whether a survey meets eligibility criteria for inclusion in the CAPS database.

- ✓ Data sets are acquired or managed by CFSAN and not intended for regulatory use
- ✓ Sample sets include food, dietary supplement, and/or cosmetic products intended for human or animal use
- ✓ Product label information is available for all samples including ingredient, nutrient, and product serving information
- ✓ Analytical data adhere to a Quality Assurance/Quality Control plan
- ✓ Product images are captured

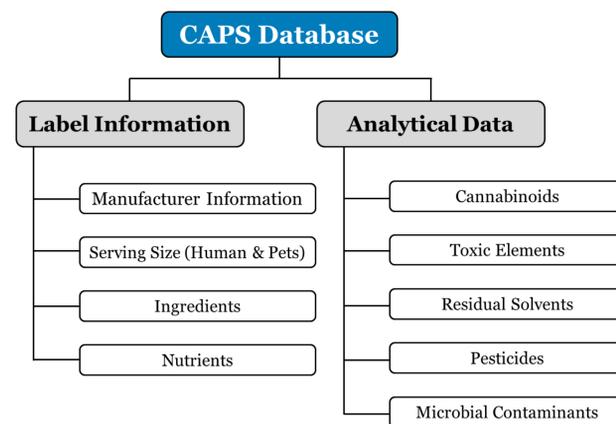
## Materials and Methods

### What is CAPS?

The CAPS database provides a platform for the management of curated analytical product survey data. Standardized data sheets are formatted to capture label information and analytical data for each product. Data are ingested into CAPS and become available to users through an interactive web-based interface <https://aps.fda.gov/>.

### Database Architecture

CFSAN's ORS/DBC and OM/DITM provide administrative oversight for CAPS. The database was developed using Microsoft SQL Server with dashboard visualizations provided by Amazon QuickSight. A schematic for the database architecture is shown in **Figure 1**.



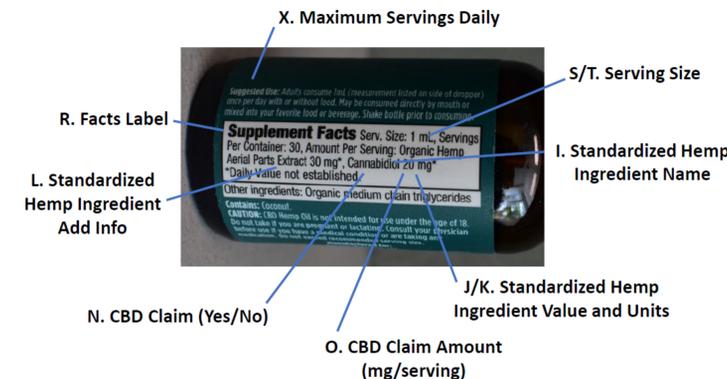
**Figure 1.** Schematic of the CAPS database architecture developed for the cannabis-derived product sampling studies.

### Data Collection and Templates

The database was designed with data integration and interoperability in mind, mimicking the construct of FDA's FoodTrak database. Data templates, guidelines, and lists of assumptions help to standardize the manual transcription of product label information.

- **Product Information Template** captures product-specific data, such as manufacturer, label claims, serving sizes, and universal product codes (UPC) (**Figure 2**). Pet products require special consideration to accommodate serving size options available for different animal body weights.
- **Ingredients Template** captures all listed ingredients, including dietary ingredients declared in the Supplement Facts panel (**Figure 3**). Listed ingredients are recorded by verbatim name and ranked in order of appearance (e.g., 1 of 8 total ingredients); those with declared amounts are reported with concentration values and units.
- **Nutrients Template** lists a standard set of 15 nutrients for all products. Declared nutrients are reported with concentration values, units, and associated percent daily value (DV) amounts.
- **Analytical Data Templates** use concentration symbols (<, =) to designate whether a value falls below, or is equal to or greater than the method detection or quantitation limit (**Figure 4**).

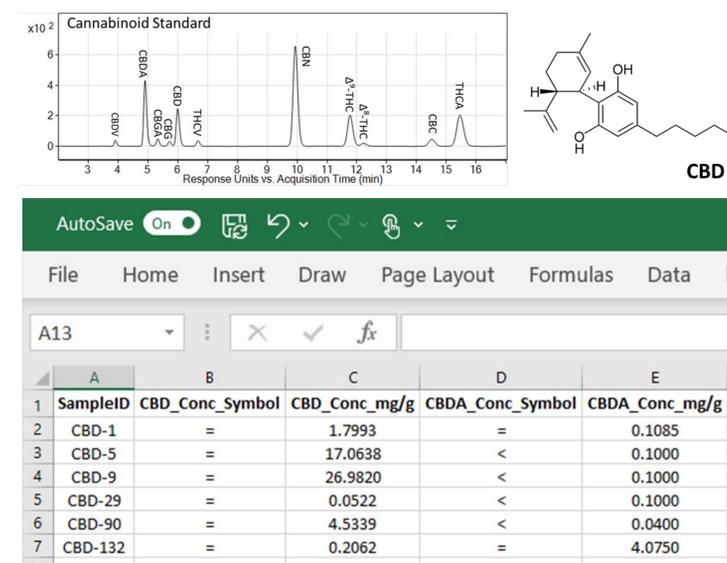
## Materials and Methods



**Figure 2.** Guidelines for capturing product-specific data, as part of the Product Information Template.



**Figure 3.** Example of (a) listed ingredients and (b) those with declared values in the Supplement Facts panel.



**Figure 4.** Analytical data template example showing cannabinoid data collected as part of the cannabis-derived product sampling study.

## Results and Discussion

### CAPS Data Resources

The CAPS user interface offers an interactive platform with data visualizations (Dashboard), product information tables (Data View), and computational calculations (Reports), creating a customized data analysis experience.

- **Dashboard** visualizations summarize data presented in the Data View (**Figure 5**). Filters can be applied to investigate targeted subsets of data. Flags are used to identify products which exceed referral level thresholds and require further evaluation.
- **Data View** data sets are selected using radio buttons for human, pet, and cosmetic products, as relevant to the sampling study. Analytical data include both static fields and computational fields (e.g., mg/serving and mg/day) for each analyte. Data View tables may be filtered, sorted, and exported.
- **Reports** page contains canned reports for computations most commonly performed for each sampling study. Label claim comparisons are made for standardized ingredients and calculations of daily contaminant exposure are possible. Additional reports list products flagged for contaminant concentrations exceeding referral level thresholds.



**Figure 5.** Dashboard screenshot showing visualizations for Count of Products by State of Origin (left) and by Product Category (right).

### Benefits of CAPS

As a benefit to the Agency, CAPS improves communication with stakeholders, provides a platform for standardized reporting, promotes data integration and interoperability, helps to inform signal detection, and supports feasibility assessments.

## Conclusion

- The CAPS database provides a standardized platform for managing product label information and analytical data generated as part of sampling studies performed or coordinated by CFSAN.
- The CAPS database provides a user-friendly, interactive tool for effectively managing big data and builds capacity for future data uses within CFSAN.