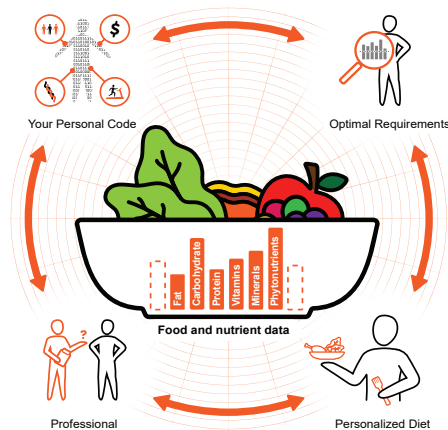


Key Insights



Food composition data inform all aspects of precision nutrition.

To identify the most critical components needed for improvement of nutrient databases, food composition data were analyzed for quality, with completeness being the most important measure, and for FAIRness, how well the data conformed with the data science criteria of findable, accessible, interoperable, and reusable (FAIR).

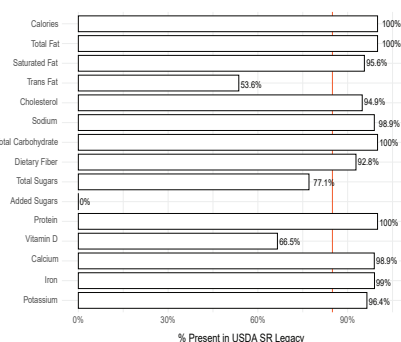
There is currently no standard, so how do we define quality?

“Minimum definition of completeness is the inclusion of all 15 Nutrition Fact Panel (NFP) nutrient measures and all 40 NASEM essential nutrient measures.”

What did we find using SR Legacy* data?

*SR Legacy provides data for 7793 foods

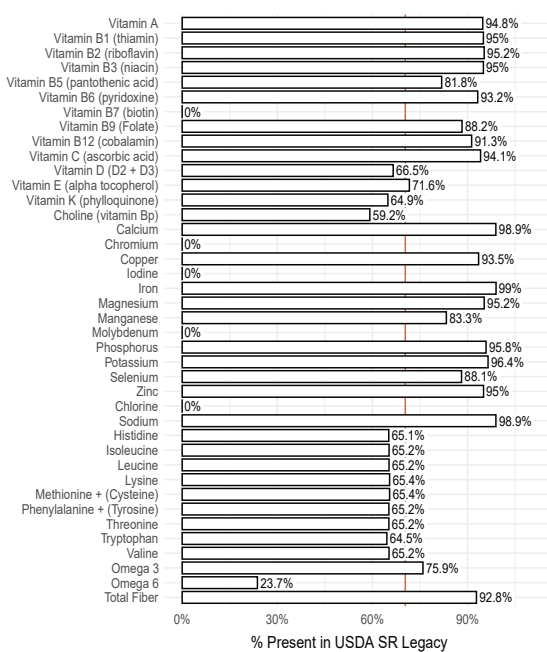
Nutrition fact panel measures



100% of foods have calorie, total carbohydrate, total fat, and protein data.

Only **67%** have **vitamin D** data; **54%** have **trans fat** data; and **0%** have **added sugars** data.

NASEM essential nutrient measures



No SR Legacy foods have **100% complete data** for NASEM essential nutrients.

Some essential nutrients are well-represented:

- iron (99%)
- calcium (99%)
- sodium (99%)

Some essential nutrients have **no data**:

- vitamin B7 (biotin)
- iodine
- chlorine/chloride
- chromium

Applying data science principles to nutrition data

“Because nutrition scientists are not data scientists, collaboration is essential for improving data quality. This includes implementing data science principles to ensure that high-quality data are available. Two key data science principles—data quality and FAIRness—are critical.”

What did we find from searching food and nutrient data sources?

Findable

How easily can a data source be found?

- Of 175 data sources, 12% were not findable.
- 32% of URL links failed at the time of collection

Accessible

How easily can data be obtained from a source?

- Of 154 findable data sources:
- 21% required **fees** or **credentials**
 - 44% were exportable in an Excel and/or CSV format
 - 47% were exportable as PDFs
 - 10% were **view-only**

Interoperable

How easily can sources be connected or harmonized?

- At present, there is no way to fully identify the true state of interoperability within the food and nutrient data landscape owing to a lack of documentation and instruction on the available linkages of food and nutrient data sources.
- USDA databases are the largest currently available globally connected network of food and nutrient data, with a network of identifiers connecting multiple other databases. But unfortunately, there is no official guide to joining them and identifier names can vary between databases.

Reusable

How easy is it to repeatedly use and access data?

- Of 137 citations, 51% did not provide metadata on either the version or the retrieval date. Because these citations do not include the version or the retrieval date, there would be no way to pinpoint what data were used.

Both data quality and data FAIRness are required for PN

“Making food composition data more complete and more FAIR will provide critical information for the implementation of PN, which can in turn have a profound positive effect on human health. PN provides an opportunity to transition away from a disease-focused model to an individualized one of

Where do we go from here?

Action steps for data suppliers and scientists:

- Collaboration between nutrition scientists and data scientists. Focusing on interoperability will foster the greatest immediate, overall improvement in database quality
- Adopt high-quality food and nutrient data standards
- PN research implements comprehensiveness and quality comparable with omics data



With PN in its infancy, this is the ideal time for nutrition scientists to recognize and elevate the importance of food and nutrient data and in doing so to set the bar high. The ambitious, comprehensive PN approach will be best positioned to develop effective diet interventions to achieve optimal health if nutrition science focuses on ensuring that food composition data are complete and FAIR.