WISEcode

Leveraging Deep Data on Foods and Food Ingredients: a WISEcode Case Study

Richard Black, PhD
Chief Science Officer

Food Forum
National Academies of Science
Washington, DC
December 5, 2025

Overview

Brief background on WISEcode

The WISEcode UPF approach to assessing processed foods (objective, data-driven, testable &modifiable)

A short demonstration of our Dashboard to access our data (with help from National Academies of Science Research Associate, Sarah Poncet

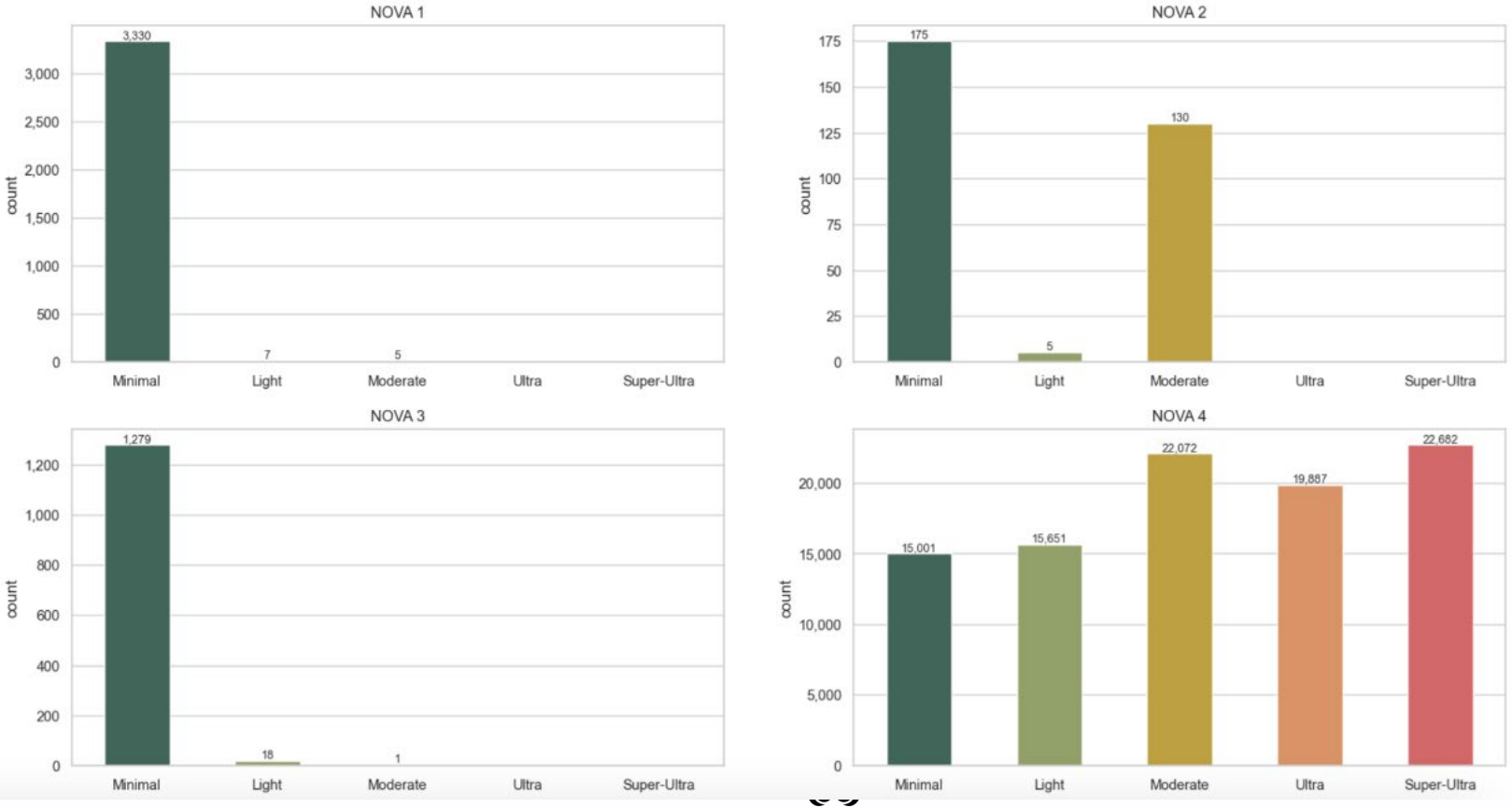
Brief Background of WISEcode

- Startup founded to develop deep nutrition data on all foods - think of USDA SR level of nutrition data depth while at the same time exceeding the breadth of the USDA Branded Foods Database - we have over 668,000 foods and over 5,000 canon ingredients (over 30,000 ingredient names).
- Dr. Matt Maxwell and Dr. Peter Frasier developed a Bayesian statistical approach to estimating the proportions of each major constituent in a food (a point estimate) and an associated error (estimate interval, generally does not exceed 5% of the point estimate).
- This enabled WISEcode to estimate the amount of essential and non-essential nutrients in a food (again, point estimate and estimate interval). However…
- We quickly learned that simply reading an ingredient was not that straightforward and we spent a couple of years learning how to do just that and then developing our Ingredient Taxonomy and Canon Ingredients.
- Each Canon Ingredient has a complete nutrient file (over 150 nutrients) as well as a variety of yes/no attributes such as: artificial; sweetener; color; flavor; preservative; emulsifier; thickener; acidulant; *etc*
- And then everyone started to get very interested not only in nutrients, but also ingredients, ingredient combinations and Ultra-processed foods (UPFs).
- Sort of the right place at the right time.

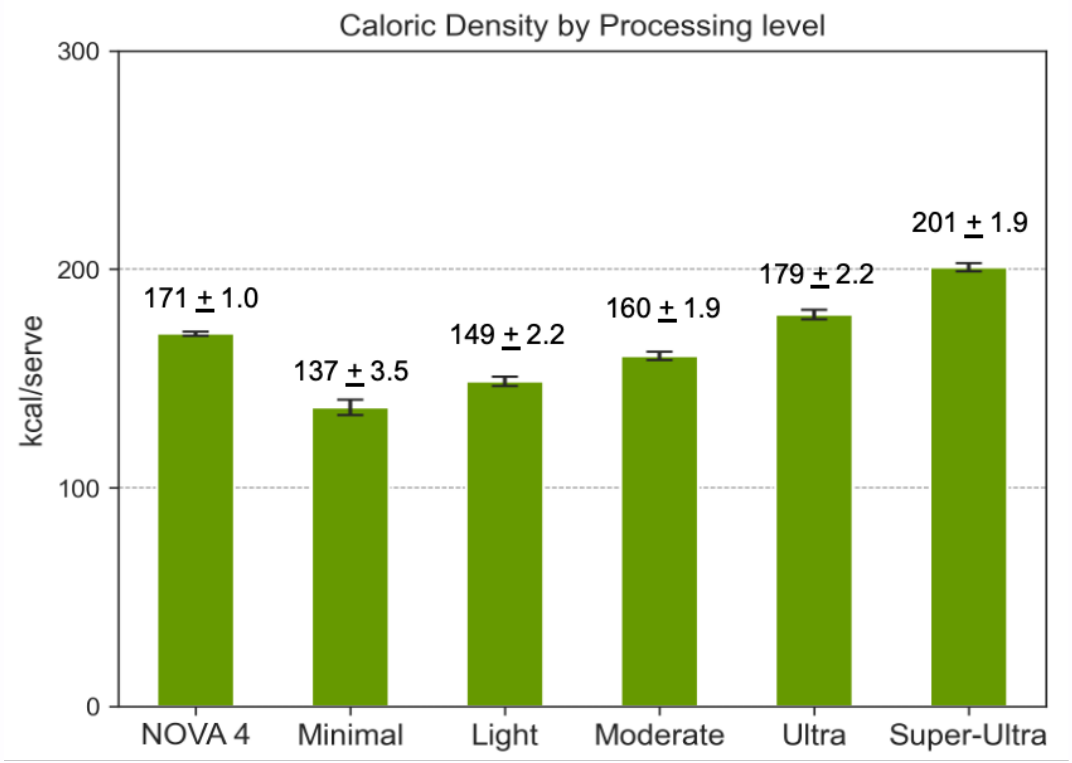
The WISEcode UPF approach to assessing processed foods (objective, data-driven, testable &modifiable)

- The Wc-UPF™ framework was designed through four iterative models, each addressing shortcomings identified in earlier versions:
  - Model 1 - Simple Count of Processed Ingredients: A basic count proved insufficient, as foods with many simple ingredients could appear more processed than foods with few but highly processed additives.
  - Model 2 - Weighted Ingredient Scoring: Ingredients were assigned to categories (1 to 4) based on degree of processing and health concern and then weighted (category minus 1). This improved differentiation but still overlooked high-sugar items.
  - Model 3 - Added Sugar Penalties: Incorporated % energy from added sugars, applying rising penalties for high-sugar foods and beverages, starting at 20% energy from added sugar and increasing by 1 for each 10% increase, capped at a penalty of 7.
  - Model 4 - Safety-Flagged Ingredients ("One-and-Done" Rule): Inclusion of any highly concerning ingredient (e.g., titanium dioxide, bromates, partially hydrogenated oils) automatically classified a food as super-ultra, regardless of score.

Wc-UPF Model4 compared with NOVA categories (100,545 foods)



WISEcode UPF distinguishes amongst foods by Caloric Density where NOVA fails to do so



The NOVA approach to UPF (NOVA 4) consistently fails to acknowledge real differences in foods

NOVA classification for Nancy's Yogurt: UPF

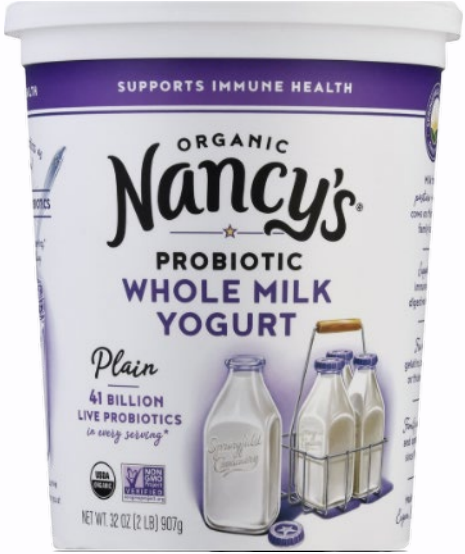
WISEcode-UPF™ for Nancy's Yogurt:
Minimal

LACTOBACILLUS CASEI:1

LACTOBACILLUS DELBRUECKII SUBSPECIES BULGARICUS:1

NONFAT MILK:1 PECTIN:1 TAPIOCA STARCH:2

WHOLE MILK:1



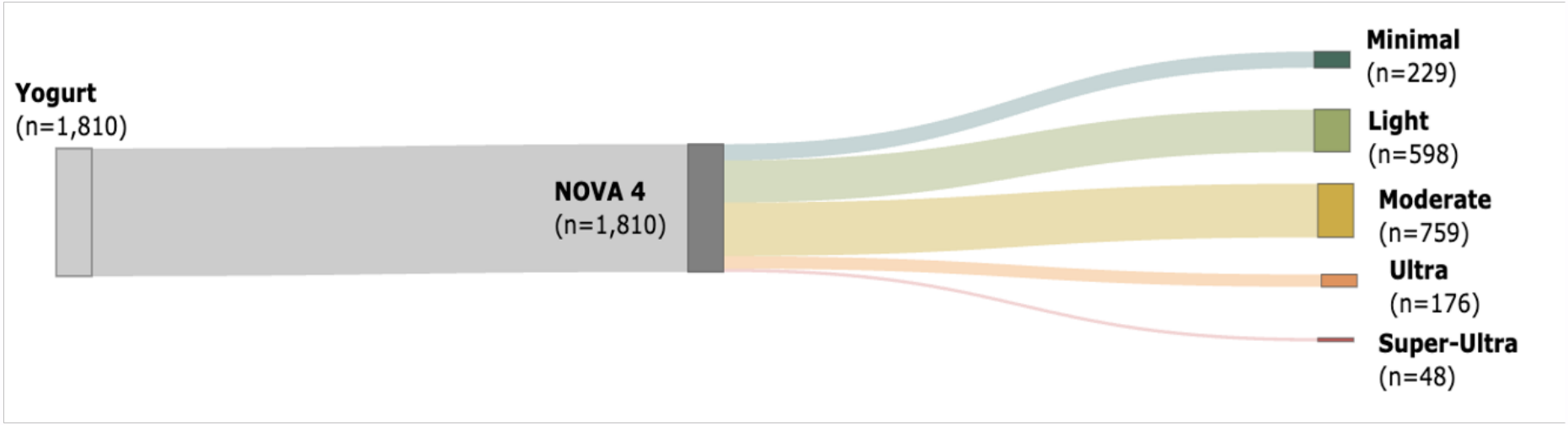
The NOVA approach to UPF (NOVA 4) consistently fails to acknowledge real differences in foods

NOVA classification for YoCrunch Yogurt: UPF

WISEcode-UPF™ for YoCrunch Yogurt: Super-Ultra

ALUMINUM LAKE FD&C BLUE NO. 1:4
ALUMINUM LAKE FD&C BLUE NO. 2:4
ALUMINUM LAKE FD&C RED NO. 40:4
ALUMINUM LAKE FD&C YELLOW NO. 5:4
ALUMINUM LAKE FD&C YELLOW NO. 6:4
ARTIFICIAL FLAVOR:2 BUTTERFAT:2 CHOCOLATE:3
CITRIC ACID:2 COCOA BUTTER:2 CORN DEXTRIN:2
CORN STARCH:2 CORN SYRUP:2
CULTURED LOWFAT MILK:1 FD&C BLUE #1:3
FD&C BLUE #2:3 FD&C RED #40:3 FD&C YELLOW #5:3
FD&C YELLOW #6:3 FRUIT JUICE:1 LACTOSE:2
M&M'S:3 MODIFIED FOOD STARCH:2 NATURAL FLAVOR:2
NONFAT MILK:1 PECTIN:1 POTASSIUM SORBATE:3
SALT:1 SODIUM CITRATE:2 SOY LECITHIN:2 SUGAR:2
VEGETABLE JUICE:1 VITAMIN A PALMITATE:1
VITAMIN D-3:1

Wc-UPF differentiates amongst foods categorized as NOVA 4



WISEcode data - Where are we now?

We now have a natural language “code builder” that accesses can be used to develop new algorithms targeting whatever is interesting to you.

Once you describe the algorithm (including weighting of various elements if desired), you can test the algorithm against a small subset of foods or against the entire data set. You can also test specific foods where you have a strong expectation of a result, or which you believe might represent outlet cases.

Our goal is not to define the processing level of foods, though we have done that. Rather, we want to have others leverage our data (free access for those in academe or at a non-profit) and take this further than the few of us at WISEcode can currently imagine.

Food data may not be sexy, but it is at the very core of understanding how diet impacts health. You just need to dive in.