

### Reliability and Validity of Assigning Ultra-Processed Food Categories to 24-Hour Dietary Recall Data Collected Using the Nutrition Data System for Research (NDS-R)

Nadia Markie Sneed,<sup>1</sup> Somto Ukwuani,<sup>2</sup> Evan Sommer,<sup>3</sup> Lauren Samuels,<sup>3</sup> Kimberly Truesdale,<sup>4</sup> Donna Matheson,<sup>5</sup> Tracy Noerper,<sup>6</sup> Shari Barkin,<sup>3</sup> and William Heerman<sup>3</sup>

<sup>1</sup>Vanderbilt University; <sup>2</sup>Vanderbilt University School of Medicine; <sup>3</sup>Vanderbilt University Medical Center; <sup>4</sup>University of North Carolina at Chapel Hill; <sup>5</sup>Stanford University; and <sup>6</sup>Lipscomb University

**Objectives:** Ultra-processed foods (UPFs) are markers of diet quality and high UPF consumption is associated with chronic disease risk. Methods to accurately categorize the degree of food processing are missing from dietary analysis programs. This study evaluated the reliability and validity of assigning the NOVA classification for UPFs to individual foods collected via 24-hour diet recalls.

**Methods:** A secondary analysis of 24-hour dietary recalls was conducted from a randomized controlled trial designed to prevent childhood obesity among low-income and minority children. Nutrition Data System for Research (NDS-R) software was used to collect 2 to 3 dietary recalls from caregivers of 610 children 3–5 years old annually over 4 years. Trained and certified coder pairs independently categorized foods into one of four NOVA classifications (minimally

processed, processed culinary ingredients, processed, and ultra-processed). The study team adjudicated any coding discrepancies. Inter-rater reliability was assessed by percent concordance between coder pairs and by Cohen's kappa coefficient. Construct validity was evaluated by comparing macronutrient estimates across NOVA classifications.

**Results:** In 5,546 valid days of 24-hour diet recalls from 610 children, there were 3,100 unique foods in the dataset, one of which was not possible to categorize. NOVA classifications for the 3,099 remaining foods were unprocessed/minimally processed (18%); processed culinary ingredients (0.4%); processed (15%); and ultra-processed (67%). Coder concordance was 88.3% with a kappa coefficient of 0.75 ( $P < .0001$ ). Descriptive comparisons of macronutrient content were consistent with expectations. On average, over the 5,546 daily recalls, UPFs made up 62% (SD 19) of the day's calories, but a higher percentage of the day's added sugar (94%; SD 16) and a lower percentage of the day's protein (47%; SD 24). Minimally processed foods made up 30% (SD 17) of the day's calories, but a lower percentage of the day's added sugar (1%; SD 8) and a higher percentage of the day's protein (43%; SD 24).

**Conclusions:** This method of applying the NOVA classification to 24-hour dietary recalls using NDS-R software had high inter-rater reliability and good construct validity for identifying individual dietary intake patterns of UPFs.

**Funding Sources:** NIH, AHRQ.