

Retest decision-making using SAS for analytical laboratory data for vitamins and minerals in adult multivitamin/mineral supplements

Joel M Palachuvattil¹, Karen W Andrews¹, Phuong-Tan Dang¹, Pavel Gusev¹, Sushma Savarala¹, Yukun Wu¹, Wendell Shirley¹, Pamela R Pehrsson¹, Johanna T Dwyer²

¹Nutrient Data Laboratory, Beltsville Human Nutrition Research Center, US Department of Agriculture (USDA), Beltsville MD
²Office of Dietary Supplements (ODS), National Institutes of Health (NIH), Bethesda MD



Background

The Nutrient Data Laboratory (NDL), Beltsville Human Nutrition Research Center (BHNRC), Agricultural Research Service (ARS) at USDA, in collaboration with the Office of Dietary Supplements, National Institutes of Health (ODS/NIH) and other federal agencies has developed a Dietary Supplement Ingredient Database (DSID; <http://dsid.usda.nih.gov>) to evaluate levels of ingredients in dietary supplement products. The DSID is funded in large part by the Office of Dietary Supplements. The goals for this project are:

- To develop reliable estimates, including variability information for nutrients and other bioactive components in DS products
- To support improved dietary intake assessments in research by providing analytical estimates of the ingredient content of marketed DSs
- To release and maintain a publicly available on-line composition database for DSs

Priority dietary supplement product categories and ingredients are determined by a DSID Working Group with members from the collaborating agencies listed above. DSID provides researchers with analytical estimates of nutrient content for adult and children's multivitamin/mineral (MVM) dietary supplements. Results from a study of adult multivitamin/mineral (MVM) products purchased in 2006-07 were released in the 2009 DSID-1 release, and updated in the 2012 DSID-2 release. A follow-up study of adult MVMs is in progress.

Objective

Improve the process of evaluating analytical lab data for retesting and acceptance of final data.

Description

124 adult MVMs were purchased in multiple lots, and samples from each lot were sent in batches of 18-20 samples to laboratories for analysis of 21 ingredients. A typical batch consisted of:

- 1 National Institute for Standards and Technology (NIST) Standard Reference Material for multivitamin/mineral 3280 (SRM)
- 2 in-house control materials, which are single lots purchased in bulk to ensure homogeneity
- 16 product samples, of which one is sent as a lab-blinded duplicate (two sets of 30 units for the same MVM product with different test sample IDs)

Labs reported the detected amount of an ingredient per gram and these data were compared to labeled levels after adjustments based on sample weights.

After evaluating the quality control (QC) results (NIST SRM, in-house controls, and duplicates), the sample results were evaluated based upon comparison to label level and average results for all samples of each supplement. Samples were flagged for retesting when the percent difference from label was higher or lower than expected, or when variability among lots was high, and when batch QC results suggested retesting.

While all results were reviewed manually, an automated process was established to aid in the review of results. This process, implemented using SAS, used an algorithm outlined in Figure 1. This algorithm decided whether a sample should be retested by taking into consideration the following factors:

- percent difference from label level
- percent difference from lowest and highest values for supplement
- percent difference from supplement mean (n>2)

This algorithm was tested with most of the results for 19 ingredients. Control results were excluded from this review. Also excluded were products with packs containing two or more types of tablets, for which additional modifications would be required.

Figure 1. Algorithm for automated review of test sample results

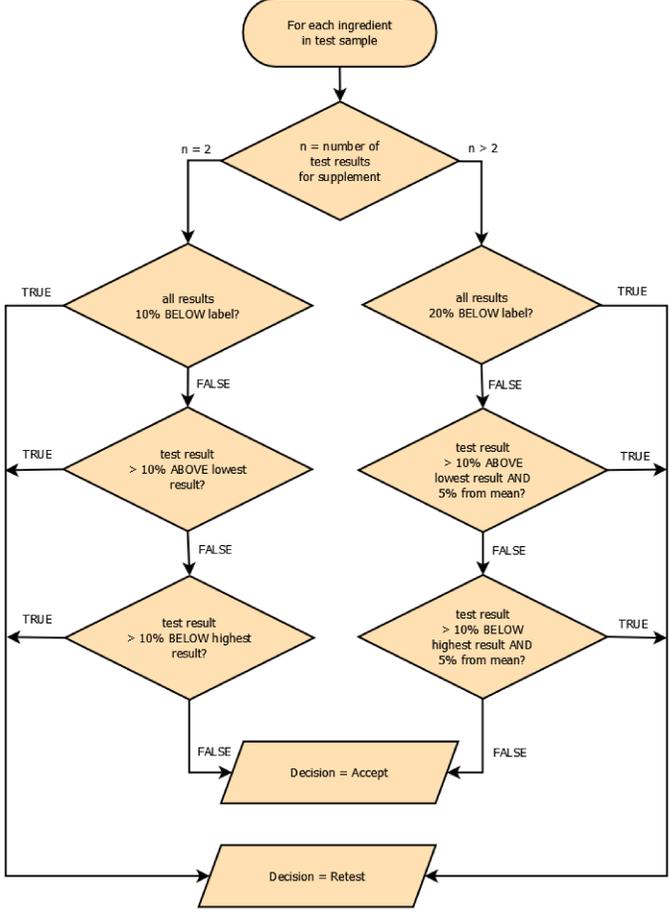


Table 1. Comparison of manual and automated review decisions

Dietary Ingredient	% Agree	% Disagree	n	κ	% Retested
Calcium	92	8	276	0.6493	16
Chromium	82	18	246	0.5050	22
Copper	86	14	265	0.5013	22
Folic Acid	82	18	324	0.5303	24
Iodine	78	22	213	0.5153	27
Iron	85	15	140	0.4806	30
Magnesium	86	14	249	0.5119	23
Manganese	85	15	251	0.3601	9
Niacin	86	14	317	0.4790	17
Phosphorus	88	12	125	0.4150	19
Potassium	92	8	112	0.6466	17
Selenium	84	16	246	0.5730	21
Vitamin C	90	10	324	0.6215	19
Vitamin E	89	11	292	0.4114	15
Zinc	91	9	277	0.5773	13

Table 2. Comparison of manual and automated review decisions for Iodine content in three samples of two MVM supplements

Supplement	Sample	Lab Result mcg/g	% diff from label	% diff from lowest result	% diff from mean	% diff from highest result	QC concern	Manual review decision	SAS algorithm decision
A	1	119	0.7	0	-12	-22	Yes	Retest	Retest
A	2	152	29	28	13	0	No	Retest	Retest
A	3	133	13	12	-1	-13	Yes	Accept	Accept
B	1	144.0	48	48.9	26.8	0.0	Yes	Retest	Retest
B	2	100.0	16	3.4	-11.9	-30.6	No	Accept	Retest
B	3	96.7	10	0.0	-14.9	-32.8	No	Accept	Retest

Discussion

For manual review, each test sample had one of three outcomes: Accept (no need to retest), Retest, and Discuss (the result was evaluated with a senior team member). For the SAS program, the possible outcomes were Accept and Retest. To evaluate the algorithm, the manual and automated results were compared for each test sample. The decisions were said to Agree if both accepted the results or both did not accept the results. The decisions were said to Disagree if one party accepted the result and the other party did not accept the result.

Table 1 shows the percentage agreement by ingredient for the number of test results (n), Cohen's kappa coefficient (κ) to measure inter-rater agreement by ingredient, and the percentage of samples actually retested. Calcium, potassium, zinc and vitamin C all had at least 90 percent agreement in decisions. Iodine results showed the most disagreement by percentage. By the standards proposed by Landis and Koch for the strength of agreement for κ, calcium, potassium and Vitamin E show substantial agreement, manganese shows fair agreement, and all other show moderate agreement.

The samples in Table 2 are results for iodine in three different lots (samples 1, 2, 3) for two different products. For Supplement A, the manual review and the SAS algorithm reached the same conclusions for retesting. There was a QC issue with two of the samples, as in both cases, one of three control results for each batch was outside of the reference range. However, other QC results for the batches were satisfactory.

There are two major reasons for disagreement between the algorithm and the manual process. First, there are instances where the algorithm will flag all supplement results for retesting, when a manual review of the results suggests only one sample should be retested. In Table 2, the mean lab result for all three samples of Supplement B is 114. Discarding the Sample 1 result, the mean of the two remaining samples is 98, and the algorithm would no longer flag samples 2 and 3 for retesting.

Secondly, the algorithm does not factor in the QC results. A high or low bias for an ingredient in a batch suggests a greater need for retests of results in that batch. The algorithm is currently not designed to account for this.

Conclusion

Improvements in storing and processing laboratory analytical data have been made to foster a consistent process for reviewing data, and considering additional factors for making retest decisions.

Automated review is clearly useful for DSID studies, with further refinement. Goals for refinement include:

- Processing results for Vitamin A, which requires combining separate assays results for retinol and beta-carotene content
- Processing results for packs
- Adjusting the decision parameters to factor in QC results
- Customize parameters by ingredient to reflect understanding of the performance of different analytical methods.

With refinement, this process can facilitate evaluating more data more efficiently in future DSID studies, and comparisons to earlier DSID study results to examine trends in DS content over time.

Other DSID Projects

DSID-3

By 2015, DSID-3 will be released. In addition to data from the original adult and children's MVM studies, it will include data from studies evaluating the levels of omega-3 fatty acids in fish, plant and fish/plant blend DSs and levels of vitamins and minerals in over-the-counter prenatal MVM products.

Botanical DSID Studies

Green tea and flavonoid-containing DSs will be the first botanical supplements to be analyzed for the DSID. Green tea DSs are among the most common botanicals purchased in the U.S. and they contain flavan-3-ols, including epicatechin, epigallocatechin, epicatechingallate, and epigallocatechingallate (EGCG), as well as caffeine. Commercial green tea dietary supplements may be dried leaves or extracts and may be chemically standardized to levels of total polyphenols, total catechins, or EGCG. In addition, many botanical DSs contain flavonoids. Flavonoids are divided into subclasses including flavonols, flavones, flavanones, flavan-3-ols, anthocyanidins, and isoflavones. The DSs most likely to contain high levels of flavonoids contain plant material or extracts from green tea, *Ginkgo biloba*, *Echinacea spp*, red clover, berries, wine, cocoa, citrus and soy.

A pilot study examining catechins and caffeine in green tea DSs is currently underway, and a pilot study examining flavonoids in botanical DSs is currently planned. The purpose of both studies is to assess methods of analysis by testing representative and top-selling products, and to identify quantitative issues in extracts and mixed herbal blends and for various label formats and ingredient levels. These studies will be carried out in consultation and collaboration with Food Composition and Methods Development Laboratory, BHNRC and other contracted laboratories. Products with a single ingredient, with multiple ingredients and with labeled and unlabeled ingredient levels will be analyzed.

The botanical pilot studies are expected to be followed by national studies, with the scope and criteria defined by the results of the pilot studies.

Reference: Landis, J.R.; & Koch, G.G. (1977). "The measurement of observer agreement for categorical data". *Biometrics* 33 (1): 159–174.

Funding: Office of Dietary Supplements, National Institutes of Health; and Nutrient Data Laboratory, Agricultural Research Service, USDA