Calculating the Nutrient Rich Foods Index Using R

Introduction

Nutrient profiling refers to assigning some form of score to foods based on their nutrient content. It is often used for ranking foods in relation to their disease prevention or health promotion properties. There are many pros and cons of nutrient profiling; a discussion of these is beyond the scope of this post, but they are discussed in detail elsewhere (Cooper, Pelly, and Lowe 2016). The purpose of this post is to describe how to use R to calculate one such nutrient profiling score—The Nutrient Rich Foods Index (Drewnowski 2009).

Nutrient Rich Foods (NRF) Index

There are a number of different NRF indices described in the work by Drewnowski (2009). For the purposes of this post, the NRF9.3 will be used. The NRF index is based on 12 nutrients; 9 nutrients (protein, fibre, vitamin A, vitamin C, calcium, iron, vitamin E, potassium, and magnesium) that are considered positive for the promotion of health and 3 nutrients (saturated fat, total sugar, and sodium) that, in excess, are known to contribute to disease.

There are a number of steps to calculating the overall score for a given food.

1. The for each of the 9 nutrients to encourage the nutrient content of the food per 100kcal is calculated as a percentage of the recommended intake for that nutrient (ED is energy density, i.e., kcal/100g). The recommended intakes used here are the US Daily Values (DV), as used in the original development paper by Drewnowski (2009). This will give 9 values, which when added together will give the NR100kcal subscore:
$$NR100kcal = \sum_{i=1}^{9}(Nutrient_i/DV_i \times 100)/ED \times 100$$
2. The above calculation is then repeated for each of the 3 nutrients to limit, to give the LIM100kcal subscore:
$$LIM100kcal = \sum_{i=1}^{3}(Nutrient_i/DV_i \times 100)/ED \times 100$$
3. Finally the subscore from the nutrients to limit is subtracted from the subscore for the nutrients to encourage to give the overall NRF score:
$$NRFscore = NR100kcal - LIM100kcal$$

Calculating the NRF Index using R

Most food composition databases are available as .csv files, and it is relatively straightforward to calculate the NRF index in this format. However, I was recently working on a project that was being carried out in R, so it made sense to write a function that would calculate the NRF score for each food in a food composition database. Using R also makes things easy if you want to alter the NRF index at a later date. For example, you may want to use different DV values or include different nutrients as was done by Drewnowski (2009).

Importing the File

For the purposes of demonstrating the function I will use the Food and Nutrient Database for Dietary Studies (FNDDS) which is openly available online from the US Department of Agriculture. The file can be read into R using the `rio` package and the following code.

```
library(rio)
dat <- import("https://www.ars.usda.gov/ARSUserFiles/80400530/apps/2017-2018%20FNDDS%20At%20A%20Glance%
```

Preparing the Data

After exploring the file, I can see it needs some tidying before use.

```
# the first row actually contains what we want to be the column names
colnames(dat) <- dat[1, ]
dat <- dat[-1, ]

# we can view what the column names are also.
# I'll just show the first 6 here so it doesn't take up too much space
names(dat[, c(1:6)])
```

```
## [1] "Food code"                  "Main food description"
## [3] "WWEIA Category number"      "WWEIA Category description"
## [5] "Energy (kcal)"              "Protein (g)"
```

```
# the nutrient content values are defined as being character rather than numeric
# these are changed to numeric
dat[, 5:69] <- sapply(dat[, 5:69], as.numeric)

# as the NRF index involves dividing by energy content, it cannot be calculated
# for energy-free foods, so these foods are removed
dat <- dat[dat$'Energy (kcal)' > 0, ]
```

Remember from above we only need 13 values for each food—12 nutrients and the energy content. So, we'll isolate these columns first. For the function below to work, it is important that the nutrients are added in the order shown below.

```
NRFdat <- dat[, c('Protein (g)', 'Fiber, total dietary (g)',
                  'Vitamin A, RAE (mcg_RAE)', 'Vitamin C (mg)',
                  'Vitamin E (alpha-tocopherol) (mg)', 'Calcium (mg)',
                  'Iron\r\n(mg)', 'Magnesium (mg)', 'Potassium (mg)',
                  'Fatty acids, total saturated (g)', 'Sugars, total\r\n(g)',
                  'Sodium (mg)', 'Energy (kcal)')]
```

Writing the Function

The function will take the the data from `NRFdat` dataframe we've just created and use this to carry out the calculations described above. It will return 3 columns of data; the first will be the subscore for the nutrients to encourage (NR100kcal), the second will be the subscore for the nutrients to limit (LIM100kcal), and the third will be the overall NRF score. The returned data will have the same number of rows as the original dataframe; each row corresponds to an individual food in the food composition database.

```r
nrffunction <- function(NRFdat){
  # first define the daily values (DV) that the food nutrient content will be
  # compared against

  # these will be taken from:
  # Drewnowski, A. 2009. "Defining Nutrient Density: Development and Validation
  # of the Nutrient Rich Foods Index." Journal Article. J Am Coll Nutr 28 (4):
  # 421S-426S. https://doi.org/10.1080/07315724.2009.10718106.

  # their order must be as follows:
  # "Protein", "Fibre", "Vitamin A", "Vitamin C", "Vitamin E", "Calcium",
  # "Iron", "Magnesium", "Potassium", "Saturated fat", "Total sugar", "Sodium"

  DV <- as.data.frame(t(c(50, 25, 800, 60, 20, 1000, 18, 400, 3500,
                                  20, 125, 2400)))

  # create an empty list to store, for each nutrient, the quantity of the
  # nutrient per 100kcal as a percentage of the reference intake for that
  # nutrient.
  nrf <- list()

  # A "for" loop will be used calculate the values for each nutrient
  # There is a cap of 100% for each nutrient.
  for(i in 1:12){
    nrf[[i]] <-
      ifelse((NRFdat[, i]/NRFdat[,13]*100)/DV[, i]*100 <= 100,
             (NRFdat[, i]/NRFdat[,13]*100)/DV[, i]*100, 100)
  }

  # convert the list to a dataframe with each column representing one of the
  # 12 nutrients in the NRF score.
  nrf <- as.data.frame(do.call("cbind", nrf))

  # subscore for nutrients to encourage
  NR100kcal <- rowSums(nrf[, c(1:9)], na.rm = TRUE)

  # subscore for nutrients to limit
  LIM100kcal <- rowSums(nrf[, c(10:12)], na.rm = TRUE)

  # overall NRF9.3 score
  NRFscore <- NR100kcal - LIM100kcal

  # create dataframe containing the above 3 values for each food
  NRFIndex <- as.data.frame(cbind(NR100kcal, LIM100kcal, NRFscore))

  return(NRFIndex)
  }
```

The function can now be used on the data above and added back to the original food composition data to give an NRF score for each food.

```r
# first apply the function to the data
NRFvalues <- nrffunction(NRFdat)
```

```r
# add the values back to the original food composition database
dat <- cbind(dat, NRFvalues)

# view the NRF scores for some of the foods
dat[c(1:5), c(2, 70, 71, 72)]
```

```
##            Main food description NR100kcal LIM100kcal NRFscore
## 2                    Milk, human  34.27449   23.23619 11.03830
## 3                      Milk, NFS  67.56349   22.26863 45.29486
## 4                    Milk, whole  50.65952   24.55222 26.10730
## 5        Milk, low sodium, whole  50.13336   23.70984 26.42352
## 6 Milk, calcium fortified, whole  50.65952   24.55222 26.10730
```

Each food now has an NRF score assigned to it and is ready for further analysis as part of the larger food composition file. The function can be updated and tailored to the specific needs of the analysis.

This post was written using R Markdown, and the .rmd file used to create it is available here:

References

Cooper, S. L., F. E. Pelly, and J. B. Lowe. 2016. "Construct and Criterion-Related Validation of Nutrient Profiling Models: A Systematic Review of the Literature." Journal Article. *Appetite* 100: 26–40. https://doi.org/10.1016/j.appet.2016.02.001.

Drewnowski, A. 2009. "Defining Nutrient Density: Development and Validation of the Nutrient Rich Foods Index." Journal Article. *J Am Coll Nutr* 28 (4): 421S–426S. https://doi.org/10.1080/07315724.2009.10718106.