In the ongoing battle over the benefits or risk of drinking coffee, a study reported in JAMA Internal Medicine has enlisted a new variable into the fray, genetics. The study makes use of genotype data on 500,000 Brits contained in the UK Biobank. The researchers first created a “caffeine metabolism” score based on SNPs, single nucleotide polymorphisms, alterations in a single base pair that had been previously identified in the scientific literature with caffeine metabolites and genes “with a plausible role in caffeine metabolism.” The result, two groups of individuals, differentiated by how quickly they metabolize caffeine. Then, using survey data on coffee consumption as well as the usual questions about weight, smoking, etc. they looked at their mortality, as reported by the National Health Service, over about a seven-year period, comparing 110,000 non-coffee drinkers to 301,000 coffee drinkers. [1]

- Coffee drinkers were more likely male, white, drinkers of alcohol and former smokers. Those drinking four cups or more a day, continued to smoke and were more likely to drink instant coffee. (Starbucks take note!) Those drinking less were more likely to be “in excellent health,” older and with a university degree.
- Concerning all-cause mortality, coffee drinking, in a dose-dependent way was protective, compared to non-coffee drinkers, reducing deaths by 14% in those drinking 8 cups a day.
- When limited to cancer and cardiovascular deaths, coffee drinking was protective although to a lesser degree.
- Ground coffee drinkers showed the most significant effect, followed by instant and decaffeinated.
- Individuals with the genetic “profile” representing faster caffeine metabolism drank more coffee
- Irrespective of the genetic “profile” coffee conferred a survival advantage. How quickly you
metabolized, caffeine made no difference.

- The exact effect of caffeine by itself seems problematic since the same trends in reducing mortality, albeit to a lesser degree, was true for those who drank decaffeinated coffee.

The study joins the growing unclear literature on the impact of coffee on our health. But it shows that our search for answers is shifting focus, from merely the amount of coffee ingested to the genetics underlying our true biologic exposure – after all, those with slower caffeine metabolisms have it hanging around for more extended periods of time. It also serves as an introduction to the term Mendelian randomization, that according to Google’s Ngram [2] appeared in about 1975, but whose use increased 63-fold by 2008.

**Mendelian randomization**

The technique was put forward as a way of improving the use of observational data in imputing cause rather than mere association. Let’s use the current study to explain the technique. The researchers want to know whether caffeine affects our mortality, but there is no randomized controlled study of identically matched individuals differentiated by their coffee drinking. In this setting, researchers may use an instrumental variable; associated with the variable thought to be the cause, but not at all related to the effect.

In this case, a genetic profile serves as the instrumental variable, having a direct effect on caffeine metabolism (and as the study showed on the amount of coffee ingested) but has no known direct effect, on mortality. And since our genes are randomly “assigned” from our parents during meiosis, we now have a way to randomize patients.

A word of caution, it is the strength of the linkage between the instrumental variable and the known variable (in this case caffeine metabolism) that increases its usefulness; and in general these genetic profiles, only account for a small percentage of effects (about 15%) so that there are still many unknown factors between association and causation. It is a more sophisticated means of randomization and stratification than what researchers already use.

Consider another example of the use of an instrumental variable, the tax on sugary beverages. We are interested in the effect of consumption of sugary drinks on health and often use taxes as the instrumental variable because more taxes lead to less consumption and have no known impact on health. If taxes were strongly linked with consumption of sugar than it would be a useful instrumental variable in searching for causation, but it isn’t, because people replace one beverage with another (juice with more sugar replacing soda). So the causal link remains problematic.

In the meantime, I’ll take a Grande Coffee Frappuccino with an extra shot.

[1] They were further stratified by coffee type with 23% drinking instant coffee and 19% drinking decaffeinated coffee.
