**Supplemental Digital Content 2.** Secondary Outcome Methods

Serum biomarkers

Blood samples (5 mL) were collected at baseline (phase 1) and after intervention (days 1 and 5) and mixed with 5 mL of EDTA to prevent coagulation. C-reactive protein, hemoglobin, interleukin-4, immunoglobulins G, E, G1, and G2a concentrations were measured as previously described8. BCM-7 was measured by electrospray ionization mass spectrometry (TripleQuad 5500, AB Sciex) with multiple reaction monitoring. Reduced glutathione levels were measured using the Glutathione Assay Kit (Sigma Aldrich).

Stool frequency and consistency

Fecal samples were collected at baseline and during each study phase (days 1, 5, 15, and 19) prior to storage at -20 °C. Fecal samples from different defecations could be put into the same collection tube but only if they were on the same day. Fecal short chain fatty acids were measured by gas chromatography-mass spectrometry8, and myeloperoxidase contents by commercially available enzyme-linked immunosorbent assay. Stool frequency and consistency were assessed via the Bristol Stool Chart in daily diaries, which subjects completed during the washout and intervention phases. Bristol Stool Score of consistency was recorded as one of the following:

1 = Separate hard lumps, like nuts (hard to pass)

2 = Sausage-shaped but lumpy

3 = Like a sausage but with cracks on the surface

4 = Like a sausage or snake, smooth and soft

5 = Soft blobs with clear-cut edges

6 = Fluffy pieces with ragged edges, a mushy stool

7 = Watery, no solid pieces. Entirely Liquid

Subtle Cognitive Impairment Test (SCIT)

Subjects indicated which of two unequal, parallel vertical lines in the target stimulus was shorter (the line on the left or that on the right). Participants indicated their choice by pressing the corresponding mouse button. The target stimuli were presented very briefly during eight different exposure durations between 16 and 128 milliseconds across 96 trials. In each trial, the presentation of the target stimulus was followed by a visual mask. The mean response time (RT; in milliseconds) and mean percentage error rate was calculated for each exposure duration. SCIT scores were recorded at baseline and after each intervention (days 1 and 15 before intervention; days 5 and 19 after intervention). The first four SCIT measures were used as the head (exposure durations between 16 and 64 milliseconds) while the tail refers to exposure durations between 80 and 128 milliseconds. SCIT response times and head and tail error rates were recorded at baseline and after each intervention (days 1 and 15 before intervention; days 5 and 19 after intervention).