J Autism Dev Disord (2016) 46:205–220
DOI 10.1007/s10803-015-2564-9

ORIGINAL PAPER

The Gluten-Free/Casein-Free Diet: A Double-Blind Challenge Trial in Children with Autism

Susan L. Hyman[1] · Patricia A. Stewart[2] · Jennifer Foley[1] · Usa Cain[1] · Robin Peck[2] · Danielle D. Morris[1] · Hongyue Wang[3] · Tristram Smith[1]

Published online: 5 September 2015

Abstract To obtain information on the safety and efficacy of the gluten-free/casein-free (GFCF) diet, we placed 14 children with autism, age 3–5 years, on the diet for 4–6 weeks and then conducted a double-blind, placebo-controlled challenge study for 12 weeks while continuing the diet, with a 12-week follow-up. Dietary challenges were delivered via weekly snacks that contained gluten, casein, gluten and casein, or placebo. With nutritional counseling, the diet was safe and well-tolerated. However, dietary challenges did not have statistically significant effects on measures of physiologic functioning, behavior problems, or autism symptoms. Although these findings must be interpreted with caution because of the small sample size, the study does not provide evidence to support general use of the GFCF diet.

Keywords Autism · Diet therapy · Gluten-free · Casein-free · Treatment outcomes

✉ Susan L. Hyman
Susan_Hyman@URMC.Rochester.edu

[1] Division of Neurodevelopmental and Behavioral Pediatrics, Department of Pediatrics, University of Rochester Medical Center, 265 Crittenden Blvd., Rochester, NY 14620, USA

[2] Department of Pediatrics and Clinical and Translational Sciences Institute, Saunders Research Building, University of Rochester Medical Center, 265 Crittenden Blvd., Rochester, NY 14642, USA

[3] Department of Biostatistics and Computational Biology, Saunders Research Building, University of Rochester Medical Center, 265 Crittenden Blvd., Rochester, NY 14642, USA

Introduction

Survey data suggests that dietary interventions are used by 15–38 % of children with autism spectrum disorder (ASD) (Interactive Autism Network 2008; Perrin et al. 2012). The most popular of these interventions is the gluten-free/casein-free (GFCF) diet (Interactive Autism Network 2008). This diet eliminates food and beverages that contain gluten (a protein found in wheat, barley, and rye) and casein (a protein found in milk and dairy products). The diet was developed to address the hypothesis that children with ASD have trouble breaking down these proteins and absorb peptides related to these compounds as a result of a leaky gut, which leads to physical discomfort and behavioral symptoms (Whiteley et al. 1999). The leaky gut was believed to allow for the entry of gluten- and casein-based peptides into the circulatory system and then into the central nervous system, where they were hypothesized to bind to opioid receptors (Horvath et al. 1999; Reichelt et al. 1991; Reichelt and Landmark 1995). Proponents of the diet propose that the resulting change in brain chemistry interferes with neural development, cognitive functioning, attention, and learning in children with ASD (Knivsberg et al. 1995).

As evidence for the leaky gut hypothesis, some studies reported abnormal levels of peptides from gluten and casein in the urine (Reichelt et al. 1994; Whiteley et al. 1999) and abnormal intestinal permeability (Horvath and Perman 2002). However, other studies have not replicated these findings (Kemperman et al. 2008; Robertson et al. 2008; Williams and Marshall 1992). Children with ASD frequently are reported to have gastrointestinal symptoms such as diarrhea or constipation, but ASD-specific gastrointestinal pathology has not been documented (Buie et al. 2010).

Despite limited evidence for the leaky gut hypothesis, narrative reports in the nonmedical literature portray dramatic success with the GFCF diet (e.g., Lewis 2011; Seroussi 2002), and some published studies corroborate these reports (reviewed by Mulloy et al. 2010). Only four published studies, however, have incorporated experimental designs to test the efficacy of the diet:

Two randomized clinical trials (RCT) reported mixed findings. Knivsberg et al. (2002) compared 10 children who were placed on the GFCF diet for 1 year to a control group of 10 children who received no dietary intervention. Relative to the control group, the GFCF group significantly improved on one outcome measure (parent-rated reduction in autistic traits) but not on three other outcome measures. In an RCT of 72 children with ASD, Whiteley et al. (2010) reported that, after 12 months, children on the GFCF diet achieved greater reductions on 5 of 11 measures of ASD, attention-deficit/hyperactivity disorder, and adaptive functioning than did children who were not on the diet. These studies had several drawbacks: First, children's adherence to the diet was not measured. Second, all outcome measures in the Knivsberg et al. (2002) RCT and most in the Whiteley et al. (2010) RCT were based on reports from children's caregivers, who were aware of whether or not their child was on the GFCF diet. Third, concomitant treatments that children may have received while enrolled in the studies were not monitored. Because of the 1-year study period, concomitant treatments could have had a substantial impact on children's outcomes. Finally, in the Whiteley et al. (2010) study, nearly a quarter of the sample was lost to follow up, yet the statistical analysis did not take into account this attrition.

The other two experimental studies in the peer reviewed literature did not detect improvement with the GFCF diet. In a crossover study with 13 children who were placed on the diet for 6 weeks and a placebo diet for 6 weeks, Elder et al. (2006) did not find changes on any outcome measure. A strength of this study was that the investigators provided food to all participants, ensuring that participants were blind to group assignment and likely to adhere to the diet or placebo. Johnson et al. (2011) found no difference on a broad range of outcome measures in an open-label, three-month RCT comparing 8 preschool children with ASD on the GFCF diet with 14 who were not on it. Nevertheless, these two studies were limited by having small sample sizes and short trials of the diet.

In one controlled study (Johnson et al. 2011), parents reported that the diet was well-tolerated and well-balanced. Further evaluation of the GFCF diet with objective measures of safety and nutritional sufficiency is needed. Many children with ASD are selective eaters (McElhanon et al. 2014), refusing all but a handful of foods or restricting themselves to foods that have a particular texture (e.g., crunchy) or color (e.g., white). Removal of gluten and casein from children's already limited diets could have a substantial impact on their nutrition. For example, removal of gluten could reduce children's intake of fortified grain products, B-vitamins, and fiber. The removal of dairy products could reduce children's intake of calcium and vitamin D, which tends to be lower than recommended even without dietary restrictions (Hyman et al. 2012; McElhanon et al. 2014). Parents might try to compensate for children's selective feeding or elimination diets by giving dietary supplements, or they might place their children on vitamin therapies also intended to change the children's behavior. These supplements often do not correct for the deficits the child exhibits or provide excess intake of nutrients above the upper limit (the intake above which side effects are likely; Stewart et al. 2015).

Overall, existing studies yield inconclusive evidence on the safety and efficacy of the GFCF diet (Millward et al. 2008; Mulloy et al. 2010). The current study aimed to address some of the design issues that compromised prior reports. First, we followed children for an extended period of time (30 weeks) while ensuring that they were in a stable educational program using similar applied behavioral analysis (ABA) methodology across participants, minimizing the potential influence of changes in concomitant treatments. Second, we employed registered dietitians who provided nutritional counseling to help families implement the diet and monitored the nutritional sufficiency of children's diets. Third, we conducted repeated checks on children's adherence to the diet. Fourth, we included physiologic outcome measures, in addition to measures of behaviors associated with ASD and behaviors that are not unique to ASD (e.g., sleep disturbance and overactivity). Fifth, the design was a double-blind placebo controlled challenge trial. Multiple measurement modalities were used (medical data from laboratory tests; ratings from parents, instructors, and research assistants; actigraphy). We established all study participants on the GFCF diet and then administered multiple, double-blind, placebo-controlled dietary challenges in a randomized, counterbalanced order. This design allowed us to examine the effects of the diet in the group of participants as a whole and in each individual participant.

Method

Design and Procedures

All study procedures were approved by the Research Subjects Review Board at the University of Rochester. We recruited children who were under 6 years old and who had received a diagnosis of Autistic Disorder, Asperger's

Disorder, or Pervasive Developmental Disorder Not Otherwise Specified based on criteria in the Diagnostic and Statistical Manual of Mental Disorders (4th ed., text revision, American Psychological Association 2000) from a large, tertiary care developmental and behavioral pediatrics clinic. This age range was selected because of the potential for therapeutic impact with younger children, the potential for similarity of the other therapies provided, and the parental control over diet.

After obtaining written, informed consent from caregivers, we conducted a *screening and initial evaluation*. At this evaluation, we confirmed the ASD diagnosis using both the Autism Diagnostic Interview (ADI-R; Rutter et al. 2003) and Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2003). We assessed cognitive functioning with the Mullen Scales of Early Learning (Mullen 1995) and adaptive functioning with the Vineland Adaptive Behavior Scales (Sparrow et al. 1984). We examined growth parameters including height and weight using World Health Organization standards (WHO Multicentre Growth Reference Study Group 2006) with three measurements, averaging the measurements if they were discrepant. Finally, we screened for medical conditions that might alter response to the GFCF diet. The screen was overseen by a board-certified developmental and behavioral pediatrician (the first author) and consisted of a parent interview, record review, physical examination, and laboratory testing in the Strong Memorial Hospital clinical laboratory for (a) tTG (tissue Transglutaminase) and IgA (Immunoglobulin A) to rule out celiac disease, (b) radioallergosorbent testing for milk, wheat, eggs and corn to identify children at risk for allergy to ingredients in the food challenges, (c) complete blood count and ferritin level to test for iron deficiency, and (d) 25-hydroxy vitamin D for vitamin D deficiency.

The remainder of the study included three phases, summarized in Fig. 1: *implementation*, *challenge*, and *maintenance*. The challenge phase incorporated a *double-blind, placebo-controlled design*. The three phases lasted a total of 30 weeks.

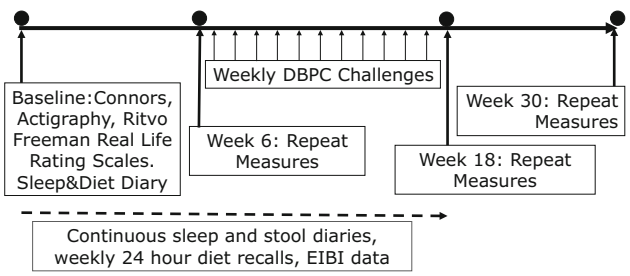

Fig. 1 Study design

Implementation Phase

We obtained baseline data on children's behavioral and nutritional status (see "Outcome Measures"). Next, we implemented the GFCF diet over 2 weeks, maintained it for at least four more weeks with weekly nutritional monitoring, and then re-assessed behavioral and nutritional status. To counsel families on implementation of the GFCF diet, the study dietitian held two sessions in the family home during the baseline period. She used written teaching materials and direct instruction to explain the diet. She also reviewed food labels, went through foods in the cupboards, and helped the family plan snacks and meals that incorporated the child's dietary preferences. She provided suggestions about food products as needed. Additionally, she guided the family on how to complete a 24-h recall of food and beverage consumed by the child; this guidance included presentation of replicas of food to illustrate portion sizes. To monitor the child's food intake and make certain the child was following the GFCF diet, the dietitian called the family weekly to obtain 24-h recalls. She used these calls to provide additional support and teaching to maintain the integrity of the diet and to assess nutrient adequacy. Families were counseled to remove casein containing products for at least 1 week and then remove gluten containing products in the next week, although most families made both changes simultaneously.

Challenge Phase

Children maintained a GFCF diet for an additional 4 weeks beyond the 2 week baseline prior to entering the challenge phase. We used a double-blind, placebo-controlled design to deliver weekly dietary challenges (see "Dietary Challenges"). Challenges occurred once per week for 12 weeks at a standard day and time determined by the child's therapy schedule. A one-week interval between challenges was chosen based on surveys indicating that 94 % of adverse reactions to gluten and casein as reported by parents resolve within a week (GFCFDiet 2001). There were four types of challenges: foods that contained gluten only, casein only, both gluten and casein, or neither (placebo). Challenges were administered in randomized, counterbalanced order. The study statistician generated the randomization sequence, which consisted of three blocks of four challenges; every block included one administration of each of the four types of challenges. The administration of three blocks of challenges is recommended in clinical studies of food allergies or intolerance in order to demonstrate the reproducibility of an adverse reaction (Metcalfe and Sampson 1990). The use of a consistent protocol across participants allowed for statistical analyses of group data obtained from the entire sample, while the

repeated challenges given to each subject in a single-subject design (N-of-1 randomized trial) allowed for inspection of data from each individual participant. (Proponents of the GFCF diet report that some individuals with ASD make dramatic improvements on the diet, whereas others do not respond at all; Seroussi 2002).

The blinded study dietitian called the family weekly during the challenge phase to provide support and guidance on implementation of the diet, assess nutritional intake, and monitor adherence. The child's behavior was observed by the research assistant, parent, and ABA therapist the day before the challenge, the day of the challenge (before and after the challenge was delivered), and 24 h after the challenge. (See "Outcome Measures") All observations occurred at the same time of day. This assessment schedule was based on clinical recommendations to monitor a child for 1–2 h after a challenge (Bock et al. 1988) and on parent reports indicating that some children with ASD show a delayed response that might not become apparent until the next day (Talk About Curing Autism 2010). If the child was not at baseline the day before the challenge, or if the child had a fever or symptoms of an intercurrent illness, the challenge was postponed until resolved. The monitoring included a visit by the research assistant to the participant's home or school, followed by counseling to the family by an investigator. If the child did not return to baseline behaviors by 24 h after a challenge, continued monitoring took place. Requiring children return to baseline before receiving an additional challenge is a standard method in clinical studies of food intolerance for ensuring that a child is neither hypo- nor hyper-sensitive to the next challenge (Metcalfe and Sampson 1990). At the end of the challenge phase, we re-assessed the child's nutritional status with a 3 day food record.

Maintenance Phase

After all 12 challenges were given, the children remained in the study for an additional 12 weeks. Families were free to maintain, modify, or abandon the GFCF diet in this phase. At the end of this period, we re-assessed the child's behavioral and nutritional status.

Participants

Eligible participants were children, aged 36–71 months at intake, with a clinical diagnosis of ASD, confirmed by the ADI-R and ADOS. To ensure stable, consistent educational services, we required children to be enrolled in a comprehensive applied behavior analysis (ABA) intervention program from one of two community agencies (see "Acknowledgments") with no changes in medication or services during the prior 4 months and no planned changes during the study period. Both agencies had well-established ABA programs founded on the UCLA/Lovaas model of early intensive ABA; instructors in both programs consistently demonstrate high fidelity to ABA procedures (Smith et al. 2015). Based on local legal mandates concerning treatment intensity, comprehensive ABA was defined as $\geq$10 h per week of one-to-one ABA intervention hours in an Individual Family Service Plan or Individualized Education Plan. All children also received speech and language therapy, special education support for developmental skills acquisition, and occupational therapy provided in collaboration with the ABA program. These services were publicly funded and provided independently of the study, based on children's individualized education plans. We planned to exclude subjects using psychotropic or sleep medication (including melatonin) because such medication might alter behaviors that also could be affected by the GFCF diet. However, no children whose caregivers expressed interest in the study were taking such medication. Other exclusion criteria included the presence of a seizure disorder, presence of a chronic illness in addition to ASD that required medical management, celiac disease, documented food allergy to wheat or milk, nutritional compromise such as iron deficiency that required treatment, and family inability to complete rating scales and assessments in English.

Based on our initial power calculation, we intended to enroll 30 children. As shown in the diagram summarizing the flow of participants through the study (Fig. 2), we screened 66 children, of whom 22 enrolled. Two of these children tested positive for tTg, requiring evaluation for celiac disease; one child left ABA; one child moved out of the area; one parent did not consistently implement the diet or record data; and two children reportedly could not be maintained on the diet because they begged for food or refused GFCF products. One child entered the study twice; this child was anemic when initially consented and did not meet diagnostic criteria for ASD when re-consented 9 months later. The remaining 14 children entered the double-blind, placebo-controlled, challenge phase of the study (described in "Design and Procedures").

Twelve of the 14 children were male. Twelve were white, 1 was African-American, and 1 was more than one race (African-American and white); all were non-Hispanic. Children's mean age at entry was 3.78 years ($SD = .60$, range = 2.96–4.97). Their cognitive skills, as measured by the Mullen (1995) Early Learning Composite standard score (SS), were varied: Three participants scored in the very low range ($SS < 49$), 8 in the low range ($SS = 50–69$), 1 below average ($SS = 70$), 1 average ($SS = 103$), and 1 above average ($SS = 120$). The Composite SS from the Vineland Adaptive Behavior Scale (Sparrow et al. 1984), obtained for 13 of the 14 participants, indicated that participants had low or moderately

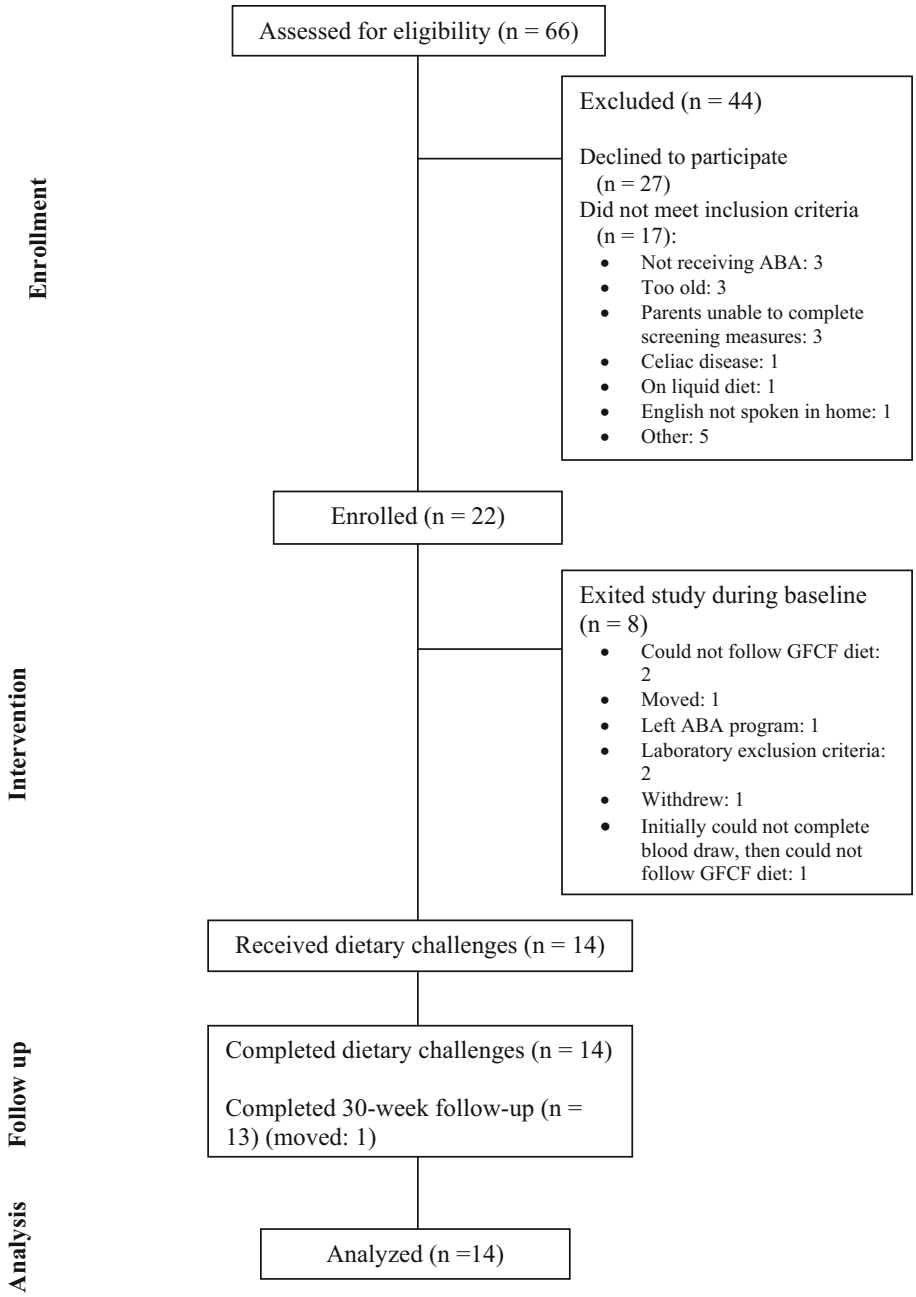

Fig. 2 Flow of participants through the study. (Diagram adapted from Shulz et al. 2010.)

low adaptive skills, *M(SD)* = 64.8(9.4), range 54–84 . The sample's Hollingshead (1975) four-point socioeconomic status (SES), a joint index of education and occupational standing, was close to the national average of 50, *M*(SD) = 51.9(10.5), range 27–66.

Dietary Challenges

Dietary challenges were delivered in the form of a snack that was individually developed for each child. At baseline, the dietitian interviewed the family about the child's taste and texture preferences. The Clinical Research Center (CRC) used this information to create snacks that contained gluten, casein, both gluten and casein, or neither (placebo). These four types of snack were indistinguishable from one another, as verified in blind taste-testing conducted in the CRC. Examples included banana bread, cookies, brownies, breakfast pastries, smoothies, puddings and egg mixtures, among others. Snacks with gluten contained 20 g of wheat flour (the calculated amount of wheat flour in two, two-inch cookies), while gluten-free snacks contained 20 g of a different type of flour (e.g., rice or tapioca flour). Snacks with casein contained 22 g of powdered cow's milk, the equivalent of one cup of reconstituted milk; snacks without

casein contained a milk substitute (soy milk). Wheat flour and cow's milk rather than pure gluten and casein were selected for the challenges in order to reflect the dietary exposures in a real diet. The doses were selected based on the serving size of typical foods consumed by children. In addition, they were intended to exceed the recommended doses in challenges for diagnosing an allergy or food intolerance (8–10 g; Sicherer 1999) and maximize the likelihood that, even if the participant consumed only part of the challenge, the dose would still exceed the minimum needed to avoid a false positive (0.15 mg/kg; Nowak-Wegrzyn et al. 2009); this increased the probability that the child would consume enough of the challenge to elicit an effect.

During the implementation phase, children were encouraged by their ABA instructors to eat a GFCF snack that resembled the dietary challenges in consistency, texture, taste, and smell. During the challenge phase, a research assistant provided the snack to the children. The child, family, ABA instructors, and research team recording behavioral data were blind to the type of challenge (gluten only, casein only, gluten + casein, or placebo). After the challenge, the research assistant brought back any portion of the snack that the child did not consume; the CRC weighed the remainder in order to estimate how much of the snack was consumed by the child.

Safety Measures

Safety measures included monitoring *growth parameters* and *blood levels* (CBC, ferritin, and vitamin D) obtained at the screening and initial evaluation. The values from the initial evaluation were used as the child's baseline; the measures were repeated at the end of each of the three phases of the study (Weeks 6, 18, and 30).

Dietary sufficiency was evaluated using a *3-Day Food Record* at baseline and at Weeks 6, 18, and 30. The parent recorded all the food and beverages consumed by the child for 3 days, including the brand, amount, and recipes for homemade foods. The dietitian provided instruction on completion of this record at baseline. Each record was analyzed for micro- and macronutrients using the Food Processor Fitness and Nutrition software (ESHA Research 2014). The 3-Day Food Record (3-DFR) is a standard technique used widely in research as well as in empirically-based clinical practice to capture recent dietary intake (Centers for Disease Control and Prevention 2014). Food Processor, originally developed in 1984 and frequently updated, is an extensively researched tool for detailed examination of nutritional intake; it uses data from sources such as the United States Food and Drug Administration, manufacturers, and restaurants to estimate intake levels for multiple nutrients.

Adverse Events

We recorded adverse events that involved behavioral deterioration (e.g., increased activity level or aggression), constitutional difficulties (e.g., diarrhea, constipation, or nutritional deficiencies), other medical concerns, or natural challenges (i.e., deviating from the GFCF diet by ingesting foods that contain gluten or casein, either administered inadvertently by an adult or obtained accidentally by the child). We recorded all adverse events that occurred during the trial period, not just the events that occurred in temporal proximity to challenges. The blinded clinician rated each event for severity and possible relatedness to the dietary challenges. All AEs were immediately reviewed by a developmental and behavioral pediatrician (SLH) blind to the content of challenges for management or referral for care. A data and safety monitoring board met routinely.

Outcome Measures

We collected outcome measures in three domains identified by proponents of the GFCF diet (e.g., Seroussi 2002) as targets of the diet: physiologic functioning, challenging behaviors (not specific to ASD), and behaviors associated with ASD.

Physiologic Functioning

Parents recorded s*tool frequency and consistency* using the Bristol Stool Scale (Lewis and Heaton 1997), which is a well-established, 7-point Likert rating scale for recording the occurrence of bowel movements and classifying the form of the feces. Parents completed the scale daily during the implementation phase and then the day before, during, and after the dietary challenge during the challenge phase.

Behaviors

Parents recorded *sleep* by completing sleep diaries daily throughout the implementation and challenge phases. The diaries included the time the child went to sleep, awoke and time and length of all night wakings. The use of sleep diaries is well-established in research on sleep disorders (Bootzin et al. 2006).

Activity and attention were measured by the Conners (1990) Abbreviated Rating Scale and actigraphy. The Conners lists 10 behaviors related to inattention, over-activity, and impulsivity. The frequency of each behavior is rated on a four-point Likert scale from 0 (not at all) to 3 (very much). The parent and ABA instructor independently completed the Conners the day before, day of, and day after each dietary challenge during the challenge phase. The

research assistant independently completed the Conners 1 day before, 2 h after the challenge, and the day after based on direct observation of the child during ABA sessions. An actigraph activity monitor, the Basic Mini-Motionlogger (Ambulatory Monitoring, Inc., Ardsley, NY), was placed on the child's wrist or waist (whichever was better tolerated by the child) for 1 h the day before each challenge and for another hour beginning 2 h after the challenge. The actigraph recorded the frequency of the participant's movements through space per 1-min interval. Variables generated were the participant's average activity level (mean number of movements per minute) and variability (standard deviation of the activity level).

Behaviors associated with ASD were measured from the Ritvo-Freeman Real Life Rating Scales (Freeman et al. 1986), based on 15 min of videotaped play with a standard set of toys and environmental arrangement. The Ritvo-Freeman contains 47 items. Each item refers to a behavior associated with ASD or a behavior that is uncharacteristic of ASD. Behaviors are rated by an observer on a four-point Likert scale from 0 (never occurs) to 3 (almost always occurs). The measure contains five scales: Sensory Motor Behaviors such as toe-walking (7 items), Social Relationships such as responding to a social bid (9 items), Affectual Reactions such as abrupt mood changes (5 items), Sensory Responses such as lining up objects into rows or gazing at twirling objects (17 items), and Language such as echolalia (10 items). Scores for ASD-related behaviors are added together, and scores for non-ASD related behaviors are subtracted from this sum; the mean score across all items is then calculated. The range of possible mean scores is −1 to 3. During the challenge phase, we administered the Ritvo-Freeman 2 h after the challenge and 1 day after the challenge. (We did not administer this instrument the day before the challenge because we sought to reduce the effects of repeated testing.) At each administration, we used one of three separate but similar sets of toys. Each set was used once during the week of each challenge; the order was counterbalanced across challenges.

The Ritvo-Freeman is an extensively studied and well-accepted measure for rating ASD-related behaviors in people from the age of 18 months to adulthood (Filipek et al. 1999); each scale has high inter-rater reliability (Freeman et al. 1986). Ritvo-Freeman ratings differentiate between children with and without ASD across a wide range of ages and developmental levels (e.g., Sevin et al. 1991). Prior to conducting assessments for the study, all new research assistants were required to establish reliability in administering (≥90 % correct implementation of the protocol) and scoring (≥80 % exact agreement with an investigator). To ensure that they maintained reliability, an investigator, who was blind to the child's dietary status, reviewed every fifth videotape for completeness and fidelity of administration as well as accuracy of scoring. The research assistants also met weekly with the first author for supervision.

Analysis

We performed analyses using SAS 9.3 (SAS Institute Inc., Cary, NC). Means and standard deviations were used to summarize study outcomes such as physiologic functioning and behavioral measurements. Repeated measure analyses of variance (SAS Proc Mixed) were performed to test the effect of dietary challenge on changes in outcome measures. In order to account for the within-subject autocorrelation caused by the repeated measurement of each child over the three blocks of challenges, the independent variable, dietary challenge, was designated as fixed effect, and the subject was designated as the random effect, with an unstructured variance covariance matrix specified (Table 3). Contrasts were further constructed to test the differences of each experimental challenge (gluten, casein, and gluten + casein) compared to the placebo challenge. Comparisons were conducted to contrast the amount of change from the day before the challenge to (a) the day of the challenge and (b) the day after the challenge. (Because the Ritvo-Freeman was obtained only the day of and day after the challenge, comparisons for this measure were performed based on the change from day of to the day after.) The Dunnett-Hsu (Hsu 1992) adjustment for multiple testing was used to maintain the family wise Type I error at 5 %. In addition, we tested the assumption of Missing Completely at Random (MCAR) and found that none of previously observed demographic variables was significantly associated with the probability of participant withdrawal. Note that the data have a monotone missing pattern (Robins et al. 1995), indicating that our assumption of MCAR is valid. The repeated measure analysis performed provides consistent estimators of regression coefficients for MCAR data (Diggle et al. 2002). Finally, we examined data on each individual participant for outcome measures that showed statistically significant ($p \leq .05$) or marginally significant ($p \leq .10$) differences between an experimental challenge and placebo.

Results

Consumption of Challenge Snacks

Children exceeded the minimum recommended dose used to elicit an effect (.15 mg/kg) in food allergy testing for all but nine challenges. These low-dose challenges occurred five times for one participant, twice for another participant, and once for each of two other participants. In addition, one

other participant moved out of the area and exited the study after the first set of four challenges. One additional participant experienced an adverse event of night-time irritability after a challenge; the participant's family asked the study team not to re-administer the same type of challenge (subsequently identified as gluten), and this participant missed one gluten challenge and two gluten + casein challenges. Of note, he experienced the same adverse events subsequent to a placebo challenge as well. One participant missed two challenges due to illness (the same participant who had five low-dose challenges). Three other participants each missed one challenge due to illness. All other participants completed challenges as scheduled.

Safety Outcomes

No clinically relevant alterations in growth parameters or lab data were observed. Because of the change in fortified foods used, one child was given an iron supplement. Another was counseled to decrease excess consumption of soy milk and bread substitutes that contained magnesium and were associated with diarrhea. Because of concerns that intake of other nutrients would be low on the diet, six children were given multivitamins; two of these children were also given calcium supplements. Two additional children were given vitamin D supplements.

Data on nutritional intake from food were obtained from 12 of the 14 participants. With nutritional counseling, the sample consumed more servings of fruits at Week 6, $M(SD) = 2.55(1.06)$, than at Baseline, $M(SD) = 1.82(.78)$, $t(10) = 4.16$, $p = .002$. There were no other statistically significant changes in consumption of fruits over the course of the study, and there were no statistically significant changes between any time points in consumption of other food groups (vegetables, grains, protein, dairy and dairy substitutes, or sweets and fats). The sample showed a significant increase in consumption of Vitamin C and fiber from baseline to Week 6: for Vitamin C (in milligrams), baseline $M(SD) = 663(440)$, Week 6 $M(SD) = 959(668)$, $t(10) = 2.53$, $p = .03$; for fiber (in grams), baseline $M(SD) = 9.85(3.86)$, Week 6 $M(SD) = 12.36(5.24)$, $t(10) = 3.10$, $p = .01$. There were no other statistically significant changes in consumption of these nutrients or others (Vitamin A and D; calcium, iron, calories, calories from fat, protein, and carbohydrates).

Adverse Events

No serious adverse events were reported during the trial. Less serious adverse events were infrequent, consisting of 1 behavioral deterioration (increased irritability reported after one challenge but not captured in the outcome measures), 7 constitutional problems (abdominal discomfort or diarrhea), 22 other medical concerns (none involving nutritional deficiencies or excesses and largely related to intercurrent illnesses), and 20 natural challenges when parents reported that children by error consumed foods containing gluten or casein. Five of the seven adverse events related to constitutional problems occurred in one participant and were similar to symptoms he had prior to the trial (the same participant who had five low-dose challenges and whose family was counseled to decrease the child's consumption of milk and bread substitutes associated with diarrhea). Apart from the natural challenges, 11 adverse events were judged by the research team at the time to be possibly or probably related to the GFCF diet or to the planned dietary challenges; of these, 9 were rated as mild and 2 as moderate in severity. Five natural challenges were associated with an increase in problem behavior observed by the parent and research assistant; these events resolved after 2–8 days. Three other natural challenges were associated with parent-reported insomnia; these events were 1–3 days in duration. Two additional challenges were followed by parent-reported increases in both problem behavior and insomnia that lasted 1–3 days. All of the families elected to continue the diet for the 12 weeks after completion of the challenges.

Outcome Assessments

Table 1 presents descriptive statistics from participants during each phase of the study: Baseline (Week 0), Implementation Phase (Weeks 1–5), end of Implementation Phase (Week 6), Challenge Phase (Weeks 7–18), and Exit (Week 30). In addition, it presents data from the day before, day of, and day after challenges during the Challenge Phase.

Table 2 presents descriptive statistics for each of the outcome measures on the day before, day of, and day after the challenge. Table 3 presents pairwise comparisons between each of the dietary challenges and placebo; comparisons are based on the change from the day before the challenge to (a) the day of the challenge and (b) the day after the challenge. (Because the Ritvo-Freeman was not administered on the day before challenges, Table 3 shows pairwise comparisons of each active challenge against placebo on the day of and the day after the challenge.) Table 2 suggests that change might have occurred on some measures but was as likely to occur with placebo challenges as with dietary challenges. For example, the sleep log (top two rows of Table 2) indicates that subjects slept somewhat less and awoke more on the day of placebo challenges than the day before but that there was little change in sleep following dietary challenges. Similarly, participants had somewhat fewer ADHD symptoms following placebo challenges, as reported by the Research

Table 1 Means and standard deviations for outcome measures in each phase of the study

Measure	All	Week				Challenge			Week 30
		0	1–5	6	7–18	Pre	Day of	Post 24 h	
Sleep log									
Minutes of sleep	630 (71.7)	618 (80.7)	639 (69.0)	631 (68.9)	627 (70.8)	633 (88.99)	605 (77.8)	638 (68.6)	641 (86.3)
# Wakings	.14 (.49)	.27 (.55)	.16 (.42)	.14 (.49)	.13 (.50)	.20 (.84)	.17 (.79)	.12 (.40)	.14 (.39)
Bristol stool[a]									
Stools per day	1.3 (.62)	1.3 (.52)	1.4 (.82)	1.4 (.62)	1.3 (.55)	1.3 (.63)	1.3 (.54)	1.3 (.58)	1.2 (.47)
Stool type	4.08 (1.40)	3.88 (1.61)	4.45 (1.54)	4.32 (1.36)	3.97 (1.33)	4.01 (1.37)	4.13 (1.40)	4.09 (1.36)	4.03 (1.25)
Conners–Rater[b]									
Research assistant	6.62 (3.83)	7.38 (4.98)	7.56 (4.62)	7.24 (3.65)	6.40 (3.45)	6.09 (2.95)	6.77 (3.73)	6.33 (3.57)	4.14 (3.62)
Teacher	8.50 (5.56)	9.22 (7.13)	9.61 (6.09)	8.67 (4.05)	8.25 (5.37)	8.42 (5.37)	8.76 (5.95)	7.53 (4.70)	7.00 (5.75)
Parent	7.36 (4.91)	10.16 (6.58)	8.10 (4.69)	7.42 (4.46)	6.98 (4.77)	6.50 (4.46)	7.57 (5.13)	6.99 (4.78)	5.32 (3.75)
Ritvo-Freeman[c]									
Sensory motor	.33 (.32)	.22 (.26)	.29 (.40)	.34 (.29)	.35 (.32)	.57 (.39)	.36 (.34)	.34 (.30)	.23 (.24)
Social relationships	−.27 (.22)	−.38 (.17)	−.29 (.23)	−.33 (.17)	−.25 (.22)	−.40 (.13)	−.24 (.24)	−.25 (.21)	−.29 (.19)
Affectual reactions	.08 (.16)	.13 (.18)	.10 (.22)	.07 (.13)	.08 (.16)	.16 (.22)	.09 (.16)	.08 (.16)	.01 (.05)
Sensory responses	.26 (.19)	.27 (.21)	.22 (.18)	.20 (.15)	.27 (.19)	.48 (.28)	.25 (.20)	.28 (.18)	.26 (.20)
Language	−.03 (.31)	−.11 (.30)	−.10 (.37)	−.03 (.31)	−.01 (.30)	−.06 (.31)	−.02 (.31)	.00 (.29)	.11 (.32)

Because actigraphy data were analyzed only for the day of challenges, the data are not included in this table

[a] Bristol Stool Scale (Lewis and Heaton 1997); [b]Conners (1990) Abbreviated Rating Scale; [c]Ritvo-Freeman Real-Life Rating Scales (Freeman et al. 1986)

Table 2 Least square means and standard deviations on the day before, day of, and day after the dietary challenge

Measure	Day before challenge				Day of challenge				Day after challenge			
	PBO	GLU	CAS	GLU + CAS	PBO	GLU	CAS	GLU + CAS	PBO	GLU	CAS	GLU + CAS
Sleep log												
Minutes of sleep	693 (56.4)	631 (67.3)	627 (76.6)	616.7 (48.3)	606 (84.2)	633 (61.5)	593 (76.8)	607 (77.0)	641 (61.3)	646 (69.1)	641 (68.2)	633 (70.7)
# Wakings	.08 (.27)	.34 (1.53)	.10 (30)	.11 (.40)	.38 (1.50)	.06 (.24)	.10 (.30)	.11 (.32)	.05 (.22)	.20 (.58)	.03 (.16)	.17 (.45)
Bristol stool[a]												
Stools per day	1.53 (.78)	1.69 (1.35)	1.62 (1.04)	1.31 (.62)	1.63 (1.13)	1.64 (.87)	1.43 (.87)	1.43 (.65)	1.75 (1.32)	1.37 (.69)	1.60 (.78)	1.44 (.91)
Stool type	3.85 (1.00)	3.86 (.84)	3.84 (1.09)	3.84 (.91)	4.08 (1.04)	3.88 (.90)	3.83 (1.42)	4.00 (1.02)	4.02 (1.09)	3.92 (1.19)	3.89 (.88)	4.23 (1.13)
Actigraph[b]												
Mean	11,234 (3640)	12,351 (3554)	9917 (3363)	10,905 (3825)	10,611 (3273)	11,251 (3491)	11,371 (3024)	11,306 (2688)	–	–	–	–
Standard deviation	5476 (785)	5641 (1107)	5566 (1865)	5719 (705)	5766 (980)	5612 (1156)	5496 (1221)	5524 (1324)	–	–	–	–
Connors–Rater[c]												
Research assistant	6.17 (3.01)	6.65 (2.38)	5.71 (2.18)	5.32 (2.13)	6.86 (2.82)	5.98 (2.81)	6.90 (1.83)	6.54 (3.27)	6.55 (2.83)	5.57 (1.61)	6.19 (2.47)	6.73 (3.24)
Teacher	8.84 (5.92)	8.50 (3.72)	8.29 (3.73)	6.89 (2.40)	9.20 (4.90)	7.95 (4.95)	8.48 (3.87)	7.76 (4.58)	7.43 (3.88)	7.76 (3.08)	7.42 (2.79)	7.19 (2.81)
Parent	5.71 (3.50)	6.10 (3.53)	6.48 (3.99)	5.79 (4.17)	7.10 (4.51)	7.02 (3.97)	6.82 (3.91)	7.70 (4.36)	7.15 (4.25)	6.75 (3.69)	6.31 (3.75)	7.15 (4.42)
Ritvo-Freeman[d]												
Sensory motor	–	–	–	–	.33 (.30)	.34 (.27)	.32 (.28)	.31 (.37)	.32 (.21)	.28 (.29)	.29 (.22)	.43 (.42)
Social relationships	–	–	–	–	−.20 (.16)	−.30 (.24)	−.30 (.21)	−.27 (.17)	−.26 (.16)	−.27 (.21)	−.28 (.18)	−.30 (.19)
Affectual reactions	–	–	–	–	.09 (.11)	.08 (.12)	.09 (.13)	.05 (.11)	.07 (.09)	.04 (.06)	.11 (.13)	.09 (.15)
Sensory responses	–	–	–	–	.30 (.19)	.25 (.14)	.26 (.19)	.21 (.14)	.32 (.18)	.29 (.17)	.27 (.16)	.26 (.17)
Language	–	–	–	–	.00 (.24)	−.07 (.31)	−.04 (.27)	−.03 (.30)	.00 (.25)	−.03 (.25)	−.03 (.25)	−.07 (.32)

CAS casein, *GLU* gluten, *PBO* placebo, *GLU + CAS* gluten + casein

[a] Bristol Stool Scale (Lewis and Heaton 1997); [b]Actigraph data analyzed only for the day before and day of challenge. Mean = mean number of movements per minute. Standard deviation = standard deviation of activity level during 1-h observation. [c]Conners (1990) Abbreviated Rating Scale; [d]Ritvo-Freeman Real-Life Rating Scales (Freeman et al. 1986), collected only the day of and day after the challenge

Table 3 Comparisons of dietary challenges to placebo. Comparisons are based on the change from the day before the challenge to (a) the day of the challenge and (b) the day after the challenge

Measure	Day of challenge									Day after challenge								
	Gluten			Casein			Gluten + Casein			Gluten			Casein			Gluten + Casein		
	t	*df*	*p*	*t*	*df*	*p*	*t*	*df*	*p*	*t*	*df*	*p*	*t*	*df*	*p*	*t*	*df*	*p*
Sleep log																		
Minutes of sleep	1.12	129	.54	−.67	129	.84	.53	129	.91	.04	129	1.00	.13	129	1.00	.29	129	.98
# Wakings	−2.26	129	.07	−1.21	129	.48	−1.17	129	.51	−.59	129	.89	−.25	129	.99	.43	129	.95
Bristol Stool[a]																		
Stools per day	−.56	128	.90	−.37	128	.97	.14	128	1.00	−1.67	132	.23	−.71	132	.82	−.14	132	1.00
Stool type	.04	91	1.00	−.01	91	1.00	.74	91	.81	−.21	90	.99	−.12	90	1.00	−.58	90	.89
Actigraphy[b]																		
Mean	−1.09	108	.56	1.54	108	.29	−.17	108	1.00	–			–			–		
Standard deviation	−.78	108	.77	−1.45	108	.34	−.88	108	.70	–			–			–		
Conners-rater[c]																		
Research assistant	−2.00	124	.12	.13	124	1.00	.15	124	1.00	−1.59	118	.27	−.07	118	1.00	.49	118	.93
Teacher	−.73	120	.81	−.48	120	.93	.24	120	.99	.66	114	.85	−.01	114	1.00	1.19	114	.49
Parent	−.28	119	.99	−.65	119	.86	−.40	119	.96	−.97	115	.65	−1.34	115	.41	−.48	115	.93
Ritvo-Freeman[d]																		
Sensory motor	.30	130	.98	−.11	130	1.00	−.71	130	.82	−1.07	126	.58	.01	126	1.00	.99	126	.63
Social relationships	−2.30	130	.06	−2.17	130	.08	−1.38	130	.38	−.72	126	.81	−1.28	126	.44	−.68	126	.84
Affectual reactions	−1.15	130	.53	−.06	130	1.00	−1.81	130	.18	−1.17	126	.51	1.76	126	.20	−.02	126	1.00
Sensory responses	−1.19	130	.49	−1.55	130	.29	−2.01	130	.12	−1.01	126	.62	−1.21	126	.48	−1.11	126	.55
Language	−2.12	130	.09	−1.22	130	.47	−1.05	130	.59	−1.12	126	.54	−.98	126	.64	−1.55	126	.29

df degrees of freedom, *p p* value with Dennett-Hsu adjustment (Hsu 1992), *t* = *t* test based on pairwise comparison with placebo

* $p > .05$ with Dunn-Bonnferoni correction for multiple comparisons (i.e., comparisons of each of the three dietary challenges with the placebo challenge)

[a] Bristol Stool Scale (Lewis and Heaton 1997); [b]Actigraph data analyzed only for the day before and day of challenge. Mean = mean number of movements per minute. Standard deviation = standard deviation of activity level during 1-h observation; [c]Conners (1990) Abbreviated Rating Scale; [d]Ritvo-Freeman Real-Life Rating Scales (Freeman et al. 1986), collected only the day of and day after the challenge; comparisons between challenges are based on comparisons of each active challenge to placebo on the day of and day after the challenge

Assistant on the Conners (seventh row of Table 2), but did not show a reduction following dietary challenges on this measure. During the Ritvo-Freeman, participants showed somewhat fewer social relationship symptoms on the day of gluten and casein challenges than on the day of placebo challenges (indicating possible improvement following consumption of gluten and casein); they also showed somewhat fewer language symptoms on the day of gluten challenges than on the day of placebo challenges (indicating possible improvement following consumption of gluten). However, none of the differences between challenges was statistically significant for any measure at any time point (Dunnett-Hsu adjusted $p < .05$), as shown in Table 3.

Individual Data

Although we did not find any statistically significant effects of active challenges compared to placebo, we did observe marginally significant effects (Dunnett-Hsu adjusted $p < .10$) for three comparisons on the day of the challenge for the Ritvo-Freeman: gluten and casein on the social relationships scale and gluten on the language scale (Table 3). Therefore, we explored the data for individual participants on these scales. The top half of Fig. 2 presents the individual data for social relationships; the bottom half of the figure shows the data for language. Each graph includes a line for each of the four experimental conditions. Higher scores indicate greater frequency of behaviors associated with ASD (i.e., worsening of symptoms). The figure shows that experimental challenges were not reliably associated with more frequent ASD behaviors. For example, ASD-associated behaviors related to social relationships for participant ST0323 (third row, second graph) may have occurred more frequently in the placebo condition than in any of the experimental challenges, although the difference was small. ASD behaviors related to social relationships for participant ST0280 (second row, fourth graph) may have occurred most frequently in response to gluten and casein challenges and least frequently in response to challenges with gluten alone. Other participants showed variable responding within and across conditions, with no clear pattern of differences among conditions. Data for the participant who had 5 challenges that may have been too low to produce an effect and who had 5 adverse events related to gastrointestinal problems (ST0021) are comparable to the data for the other participants. Data for the participant who exited the study early (ST0236), the participant whose family requested that the study team stop administering challenges that were subsequently identified as gluten (ST0068), and the participants who each missed one challenge due to illness (ST0065, ST0277, and ST0281) are also comparable to the data for the other participants. Three of these participants (ST0065, ST0277, and ST0281) and two additional participants (ST0324, and ST0404) also had missing data due to cancelled sessions; their data are similar to the data for other participants. Thus, the individual data suggest that the experimental challenges did not have a consistent effect on behavior and demonstrate that the group statistics did not obscure individual response (Fig. 3).

Discussion

Our double-blind, placebo-controlled study of 14 children on the GFCF diet did not detect an impact of dietary challenges on measures of physiologic function, behavioral disturbance (sleep disruption and overactivity), or ASD-related behaviors. One participant had reported reaction to gluten that was not captured by our data and that was also reported in response to placebo. To maximize the opportunity to observe an effect, we provided extensive nutritional counseling to establish and stabilize children on the diet over a period of 6 weeks before introducing the challenge, and delivered challenges over a subsequent 12-week period. We also included a broad range of outcome measures, obtained from diverse sources (medical data from laboratory tests, rating scales completed by several different adults, actigraphy, and direct observation). With ongoing nutritional counseling and monitoring, the diet appeared to be adequate for most participants; individualized supplementation was added for a few participants when deemed necessary by the study dietitian to address low intake of iron, calcium, or vitamin D.

Several features of our study are novel in the literature of dietary interventions used for behavior change. First, we employed registered dietitians to closely monitor children's adherence to the diet and the child's nutritional intake. Second, we carefully measured children's consumption of gluten and casein when given challenge snacks. Third, we ensured that children were receiving stable, consistent educational and behavioral services by recruiting children who were receiving ABA intervention from community agencies with a documented history of implementing ABA techniques with fidelity (Smith et al. 2015). Fourth, we used physiologic as well as behavioral measures of outcome. Fifth, the subjects, families, research assistants, teachers and all caregivers were blinded to the challenge given. Finally, we obtained ratings from multiple informants and also incorporated objective measures (laboratory tests and actigraphy).

Our findings accord with those of Elder et al. (2006) and Johnson et al. (2011), who did not find evidence of benefit from the GFCF diet. Our findings differ, however, from results presented by Knivsberg et al. (2002) and Whiteley

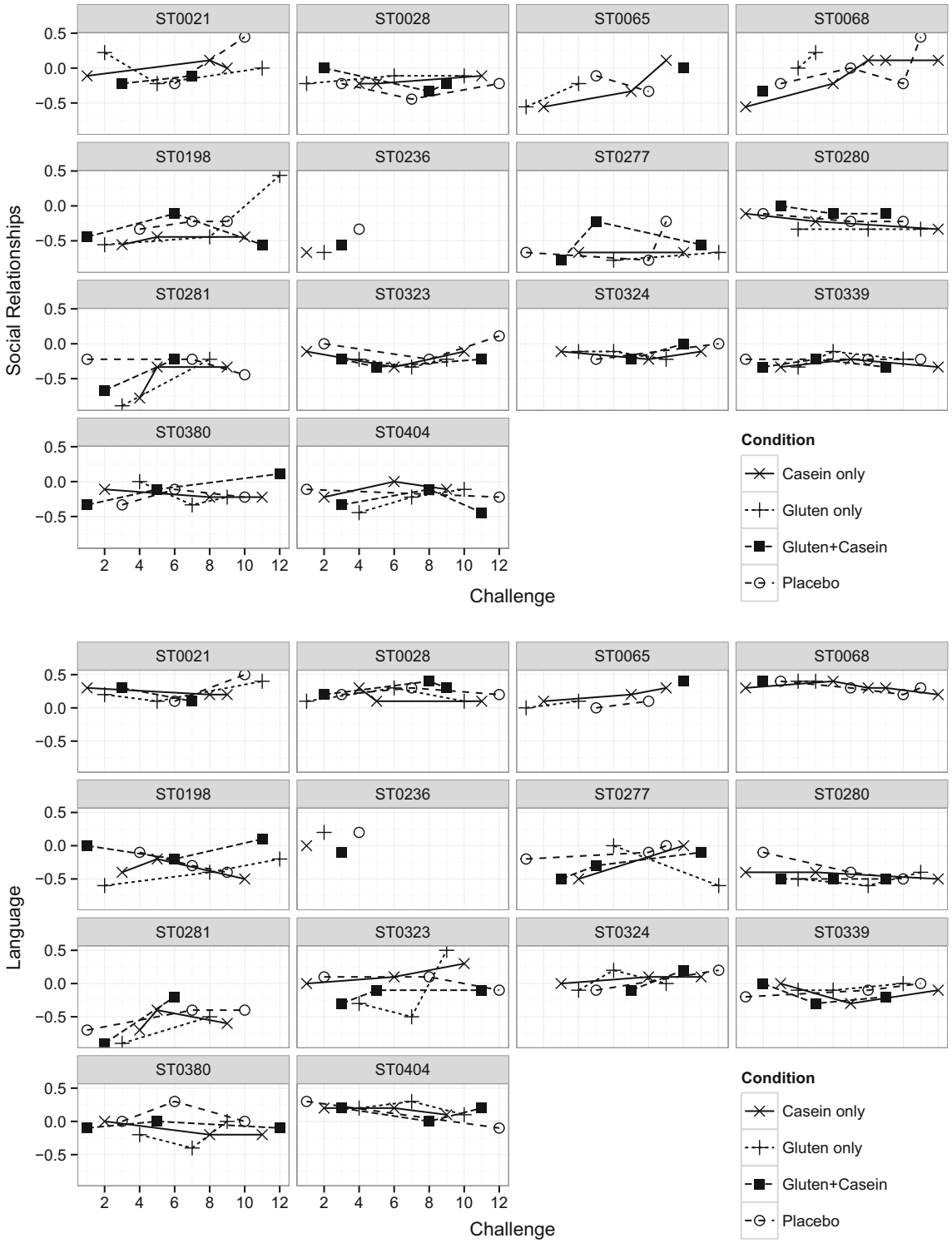

Fig. 3 Data on individual participants from the Ritvo-Freeman Real-Life Rating Scales (Freeman et al. 1986). The *top half* of the figure presents data from the social relationships scale; the *bottom half* of the figure shows data from the language scale. Each graph depicts data for one participant. Each *line* represents data from one dietary challenge condition

et al. (2010); these investigators found benefits of the GFCF diet on some (but not most) outcome measures. One possible reason for this difference is that the positive findings described by Knivsberg et al. and Whiteley et al. were obtained primarily from measures that were based on report by parents who were aware of whether or not their

children were on the GFCF diet. Another possible reason is that their findings were based on a one-year trial in which participants received a range of interventions outside the study. Concomitant treatments were not monitored.

Perhaps the most notable limitation of our study is the small sample size. Despite the popularity of the GFCF diet, recruitment was slower than we anticipated. This might have been because information became widely available in the popular media to guide families who wish to implement the diet on their own during the course of the study. Indeed, many families who expressed interest in the diet declined to enroll in the study, as shown in the subject flow diagram (Fig. 2). In addition, recent surveys suggest that fewer children receive the diet than previously reported—15–16 %, compared to prior estimates as high as 38 %; (Interactive Autism Network 2008; Perrin et al. 2012). It is unclear whether this diet is declining in popularity, whether other nutritional interventions are gaining popularity, or whether the varying estimates reflect methodological differences across studies. Also, the rigor of the diet and study protocol was a substantial barrier. Only 14 of the 22 enrolled participants and their families successfully implemented the diet and collected data, even with frequent, professional direction from the study team. Moreover, some of these 14 participants missed one or more challenges or had other missing data. The small sample size limits interpretation and generalizability of findings.

Another concern is that some proponents of the diet might regard the 4–6 week implementation phase prior to the challenges as too short for the GFCF diet to take full effect (Seroussi 2002); if so, the introduction of dietary challenges in the next 12 weeks of the study could have been premature. However, this is improbable. Although serologic and histologic improvement in patients with celiac disease may take a prolonged period, symptom improvement occurs within days for diarrhea and within a month for abdominal pain (Murray et al. 2004). Even with nonceliac gluten sensitivity, the diagnostic evaluation includes a 6 week gluten free diet (with reported improvement of symptoms within days), then a one week challenge period followed by a 1 week reversal (Catassi et al. 2015). As previously noted, food allergy symptoms emerge within 2 days of exposure, so 4–6 weeks is adequate for resolution. Given these considerations and the rapidity with which proponents report benefit from the diet, 4–6 weeks was judged to be adequate to introduce the diet.

Another possible concern is that we focused on removing gluten and casein from children's diet, but some investigators place equal importance on removing other items such as corn or soy (Matthews 2008). Additionally, we excluded children who had known gastrointestinal disorders, who might have been more likely to respond positively to dietary restriction. Children with lactose intolerance, for example, might have intestinal discomfort and diarrhea with milk potentiatng irritability. The prescreening identified two children at risk for celiac disease who were referred to a gastroenterologist for further assessment and not included in this study. No children were identified as being at risk for allergy to wheat or milk based on IgE testing to these food items. This study did not investigate the potential impact on behavior of diets that alter gut flora (Halmos et al. 2015).

We do not think that the weekly interval for challenges was too frequent, since all children returned to behavioral baseline prior to receiving the next challenge. An informal parent based survey reported on the web site www.gfcfdiet.com at the time this study was planned indicated that almost all children who had a dietary infraction while on the diet were perceived as returning to their baseline within 1 week, most within a few days. We believe the dose of dietary challenges was adequate, given that it almost always exceeded the threshold considered necessary to elicit an effect in testing for food allergies and given that parents have reported effects even from tiny doses. Our data indicates that baseline behaviors were present at 24 h after the challenge.

In sum, although our findings must be interpreted with caution because of our small sample size and the other noted study limitations, our study does not provide evidence to support general use of the GFCF diet. Additionally, although we cannot comment on the potential use of the diet for children with ASD who have gastrointestinal symptoms, our findings indicate that the diet can be safe for other children with ASD if appropriately implemented and monitored for nutritional sufficiency. The design of this trial controlled for concurrent use of therapy known to be effective so that the impact of diet was not attributed to other interventions a child might be receiving.

Acknowledgments The project described in this publication was funded by U54 MH066397 (NIH/NIMH Genotype and Phenotype of Autism, Principal Investigator: Patricia Rodier). It was also supported by the University of Rochester CTSA award UL1 RR024160 from the National Center for Research Resources and the National Center for Advancing Translational Sciences of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank the children and families who participated. We also thank Eileen Blakely, R.D., for assisting with nutritional monitoring; Emily Healy for assisting with data collection; Patricia Rodier, Ph.D. (deceased), for her leadership of the STAART Center; Caroline Magyar, Ph.D., and her staff for conducting intake assessments of participants; Christopher Stodgell, Ph.D., for directing data management of the project; and Loisa Bennetto, Ph.D., for her advice. We are grateful to the following agencies who referred participants to the study, participated in study implementation, and advised us: Stepping Stones Learning Center (Mariellen Cupini, Director; Dr. John McEachin, EIBI Consultant) and the Center for Autism and Related Disorders (Fairport, NY branch, Director: Denise Rhine; Dr. Doreen Granpeesheh, Executive Director).

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Bock, S. A., Sampson, H. A., Atkins, F. M., Zeiger, R. S., Lehrer, S., Sachs, M., & Metcalfe, D. D. (1988). Double-blind, placebo-controlled food challenge (DBPCFC) as an office procedure: a manual. *Journal of Allergy and Clinical Immunology, 82*(6), 986–997.

Bootzin, R., Buysse, D., Edinger, J., Espie, C., Lichstein, K., & Morin, C. (2006). Psychological and behavioral treatment of insomnia: Update of the recent evidence (1998–2004). *Sleep, 29*, 1398–1414.

Buie, T., Campbell, D. B., Fuchs, G. J, I. I. I., Furuta, G. T., Levy, J., Van de Water, J., & Winter, H. (2010). Evaluation, diagnosis, and treatment of gastrointestinal disorders in individuals with ASDs: A consensus report. *Pediatrics, 125*(Supplement 1), S1–S18.

Catassi, C., Elli, L., Bonaz, B., Bouma, G., Carroccio, A., Castillejo, G., & Fasano, A. (2015). Diagnosis of non-celiac gluten sensitivity (NCGS): The Salerno experts' criteria. *Nutrients, 7*(6), 4966.

Centers for Disease Control and Prevention (2014). *National Health and Nutrition Examination Survey, 2013–2014*. Retrieved from http://www.cdc.gov/nchs/nhanes/about_nhanes.htm.

Conners, C. K. (1990). *Conners' Abbreviated Symptom Questionnaire: Parent version, teacher version manual*. Toronto: Multi-Health Systems.

Diggle, P. J., Heagerty, P., Liang, K. Y., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). New York: Oxford University Press.

Elder, J. H., Shankar, M., Shuster, J., Theriaque, D., Burns, S., & Sherrill, L. (2006). The gluten-free, casein-free diet in autism: Results of a preliminary double blind clinical trial. *Journal of Autism and Developmental Disorders, 36*, 413–420.

ESHA Research (2014). *Food Processor Nutrition Analysis and Fitness Software: Full Manual.* Retrieved from http://www.esha.com/sites/default/files/documents/full_manual.pdf.

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H, Jr, Dawson, G., Gordon, B., & Volkmar, F. R. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29*, 439–484.

Freeman, B. J., Ritvo, E. R., Yokota, A., & Ritvo, A. (1986). A scale for rating symptoms of patients with the syndrome of autism in real-life settings. *Journal of the American Academy of Child and Adolescent Psychiatry, 25*, 131–136.

GFCFDiet (2001). *Parent survey GFCF diet.* Accessed November 6, 2014, at http://www.gfcfdiet.com/dietsurveysept2.htm.

Halmos, E. P., Christophersen, C. T., Bird, A. R., Shepherd, S. J., Gibson, P. R., & Muir, J. G. (2015). Diets that differ in their FODMAP content alter the colonic luminal microenvironment. *Gut, 64*(1), 93–100. doi:10.1136/gutjnl-2014-307264.

Hollingshead, A. B. (1975). *Four-factor index of social status*. Retrieved from http://elsinore.cis.yale.edu/sociology/yjs/yjs_fall_2011.pdf#page=21.

Horvath, K., Papadimitriou, J. C., Rabsztyzn, A., Drachenberg, C., & Tildon, J. T. (1999). Gastrointestinal abnormalities in children with autistic disorder. *Journal of Pediatrics, 135*, 559–563.

Horvath, K., & Perman, J. A. (2002). Autistic disorder and gastrointestinal disease. *Current Opinion in Pediatrics, 14*, 583–587.

Hsu, J. C. (1992). The factor analytic approach to simultaneous inference in the general linear model. *Journal of Computational and Graphical Statistics, 1*(2), 151–168. doi:10.1080/10618600.1992.10477011.

Hyman, S. L., Stewart, P. A., Schmidt, B., Lemcke, N., Foley, J. T., Peck, R., & Handen, B. (2012). Nutrient intake from food in children with autism. *Pediatrics, 130*(Supplement 2), S145–S153. doi:10.1542/peds.2012-0900L.

Interactive Autism Network (IAN, 2008). *IAN Research Findings: Special Diets*. Retrieved from http://www.iancommunity.org/cs/ian_treatment_reports/special_diets.

Johnson, C. R., Handen, B. L., Zimmer, M., Sacco, K., & Turner, K. (2011). Effects of gluten free/casein free diet in young children with autism: A pilot study. *Journal of Developmental and Physical Disabilities, 23*(3), 213–225. doi:10.1007/s10882-010-9217-x.

Kemperman, R., Muskiet, F., Boutier, A., Kema, I., & Muskiet, F. (2008). Normal intestinal permeability at elevated platelet serotonin levels in a subgroup of children with pervasive developmental disorders in Curacao (the Netherlands Antilles). *Journal of Autism and Developmental Disorders, 38*, 401–406. doi:10.1007/s10803-007-0399-8.

Knivsberg, A. M., Reichelt, K. L., Høien, T., & Nødland, M. (2002). A randomised, controlled study of dietary intervention in autistic syndromes. *Nutritional Neuroscience, 5*(4), 251–261.

Knivsberg, A. M., Reichelt, K. L., Nødland, M., & Høien, T. (1995). Autistic syndromes and diet: A follow-up study. *Scandinavian Journal of Educational Research, 39*(3), 222–236.

Lewis, L. (2011). *Special diets for special kids, Volumes 1 and 2 combined: Over 200 revised and new gluten-free casein-free recipes, plus research on the positive effects for children with autism, ADHD, allergies, celiac disease, and more!*. Arlington, TX: Future Horizons.

Lewis, S. J., & Heaton, K. W. (1997). Stool form scale as a useful guide to intestinal transit time. *Scandinavian Journal of Gastroenterology, 32*(9), 920–924. doi:10.3109/00365529709011203.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2003). *Autism diagnostic observation schedule*. Los Angeles, CA: Western Psychological Services.

Matthews, J. (2008). *Nourishing hope for children with autism: Nutrition and diet guide for healing our children*. San Francisco: Healthful Living Media.

McElhanon, B. O., McCracken, C., Karpen, S., & Sharp, W. G. (2014). Gastrointestinal symptoms in autism spectrum disorder: A meta-analysis. *Pediatrics, 133*(5), 872–883.

Metcalfe, D. D., & Sampson, H. A. (1990). Workshop on experimental methodology for clinical studies of adverse reactions to foods and food additives. *Journal of Allergy and Clinical Immunology, 86*(3), 421–442.

Millward, C., Ferriter, M., Calver, S. J., & Connell-Jones, G. G. (2008). Gluten- and casein-free diets for autistic spectrum disorder. *Cochrane Database of Systematic Reviews,*. doi:10.1002/14651858.CD003498.pub3.

Mullen, E. M. (1995). *Mullen Scales of early learning*. Circle Pines, MN: American Guidance Service.

Mulloy, A., Lang, R., O'Reilly, M., Sigafoos, J., Lancioni, G., & Rispoli, M. (2010). Gluten-free and casein-free diets in the treatment of autism spectrum disorders: A systematic review. *Research in Autism Spectrum Disorders, 4*(3), 328–339.

Murray, J. A., Watson, T., Clearman, B., & Mitros, F. (2004). Effect of a gluten-free diet on gastrointestinal symptoms in celiac disease. *American Journal of Clinical Nutrition, 79*(4), 669–673.

Nowak-Wegrzyn, A., Assa'ad, A. H., Bahna, S. L., Bock, S. A., Sicherer, S. H., & Teuber, S. S. (2009). Work Group report: Oral food challenge testing. *Journal of Allergy and Clinical Immunology, 123*(6 Supplement), S365–S383. doi:10.1016/j.jaci.2009.03.042.

Perrin, J. M., Coury, D. L., Hyman, S. L., Cole, L., Reynolds, A. M., & Clemons, T. (2012). Complementary and alternative medicine

use in a large pediatric autism sample. *Pediatrics, 130*(Supplement 2), S77–S82. doi:10.1542/peds.2012-0900E.

Reichelt, K. L., Knivsberg, A., Lind, G., & Nødland, M. (1991). Probable etiology and possible treatment of childhood autism. *Brain Dysfunction, 4*, 308–319.

Reichelt, K. L., Knivsberg, A., Nødland, M., & Lind, G. (1994). Nature and consequences of hyperpeptiduria and bovine casomorphins found in autistic syndromes. *Developmental Brain Dysfunction, 7*, 71–85.

Reichelt, K. L., & Landmark, J. (1995). Specific IgA antibody increases in schizophrenia. *Journal of Biological Psychiatry, 37*, 410–413.

Robertson, M., Sigalet, D., Holst, J., Meddings, J., Wood, J., & Sharkey, K. (2008). Intestinal permeability and glucagon-like peptide-2 in children with autism: A controlled pilot study. *Journal of Autism and Developmental Disorders, 38*, 1066–1071. doi:10.1007/s10803-007-0482-1

Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association, 90*, 106–121.

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism diagnostic interview-revised*. Los Angeles, CA: Western Psychological Services.

Seroussi, K. (2002). *Unraveling the mystery of autism and pervasive developmental disorder: A mother's story of research & recovery*. New York, NY: Crown.

Sevin, J. A., Matson, J. L., Coe, D. A., Fee, V. E., & Sevin, B. M. (1991). A comparison and evaluation of three commonly used autism scales. *Journal of Autism and Developmental Disorder, 21*, 417–432.

Sicherer, S. H. (1999). Food allergy: When and how to perform oral food challenges. *Pediatric Allergy and Immunology, 10*(4), 226–234.

Smith, T., Klorman, R., & Mruzek, D. W. (2015). Predicting outcome of community-based early intensive behavioral intervention for children with autism. *Journal of Abnormal Child Psychology,*. doi:10.1007/s10802-015-0002-2.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales* (Survey ed.). Circle Pines, MN: American Guidance Service.

Stewart, P. A., Hyman, S. L., Schmidt, B. L., Macklin, E. A., Reynolds, A., Johnson, C. R., et al. (2015). Dietary supplementation in children with autism spectrum disorders: Common, insufficient and excessive. *Journal of the Academy of Nutrition and Dietetics, 115*, 1237–1248. doi:10.1016/j.jand.2015.03.026.

Talk About Curing Autism (2010). *Diet infringement or infraction help*. https://www.tacanow.org/family-resources/diet-infringement-or-infraction-help/. Accessed 21 Aug 2015.

Whiteley, P., Haracopos, D., Knivsberg, A. M., Reichelt, K. L., Parlar, S., Jacobsen, J., & Shattock, P. (2010). The ScanBrit randomised, controlled, single-blind study of a gluten- and casein-free dietary intervention for children with autism spectrum disorders. *Nutritional Neuroscience, 13*(2), 87–100. doi:10.1179/147683010x12611460763922.

Whiteley, P., Rodgers, J., Savery, D., & Shattock, P. (1999). A gluten-free diet as an intervention for autism and associated spectrum disorders: Preliminary findings. *Autism, 3*(1), 45–65. doi:10.1177/1362361399003001005.

WHO Multicentre Growth Reference Study Group. (2006). *WHO child growth standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development*. Geneva, World Health Organization, 2006. Retrieved from http://www.who.int/childgrowth/standards/Technical_report.pdf?ua=1.

Williams, K. M., & Marshall, T. (1992). Urinary protein patterns in autism as revealed by high resolution two-dimensional electrophoresis. *Biochemical Society Transactions, 20*, 189S.