Systematic Pediatric Assessment of Rome Criteria (SPARC)

Principal Investigator: William Bennett MD

IRB Protocol # 1811325201

NCT04773158

April 1, 2019

Sample Size and Power Estimates. We calculated our sample size using the binary primary outcome (whether there was resolution of symptoms of a FGID by Rome IV criteria in children and adolescents or whether parental reassurance was achieved for infants). We wish to detect a 15% difference between the control and intervention clinics in the rate of the primary outcome. Since each FGID has a different estimated background rate of improvement without intervention, we estimated an overall weighted average improvement rate. We also took into consideration of differences of prevalence of different FGID diagnosis. Therefore, based on the data presented in Table 3, we estimated an overall improvement rate of 30% for the control clinics. Though we do not anticipate much variability in this rate at our six study clinics, we will use six different rates in a wide range, 15%, 20%, 25%, 30%, 35% and 40%, at baseline in order to estimate intra-clinic correlation. This translates to an intra-class correlation coefficient 0.007. However, for a conservative estimation of sample size, we will use intra-class correlation 0.014, two times the intra-class correlation we estimated.

Table 3. Data Utilized for Sample Size Calculations

FGID Diagnosis	Estimated Baseline Improvement in 3 months	Prevalence	Unique CHICA patients/yr in age group	Minimum Time Required
Infant regurgitation (0-12m)	25%	10%	3172	10 months
Infant dyschezia (0-9m)	25%	5%	2640	26 months
Infant colic (0-5m)	48%[77]	10%	1584	26 months
Cyclic vomiting syndrome (3m-18y)	70%[78]	5%	22544	4 months
Functional nausea/vomiting (8-18y)	50%	5%	9702	8 months
Functional dyspepsia (8-18y)	25%	10%	9702	4 months
Irritable bowel syndrome (8-18y)	37.5%[79]	15%[80]	9702	4 months
Abdominal migraine (8-18y)	20%	2%[81]	9702	16 months
Functional abdominal pain - NOS	20%	15%	9702	4 months
Functional constipation (0-18y)	20%	20%	23336	3 months

Using a chi-square test and setting alpha at 0.05, the probability of detecting a statistically significant difference in the proportion of children with symptoms of any FGID with improvement in the primary outcome at 3 months from baseline between the intervention and the control group is 90%, with an effective sample size of 163 per group. Hence, the total effective sample size (n*) is 326.

The unit of randomization is the clinic, but the unit of analysis is the individual patient. Since patients are nested within practice, responses from the patients within practice are likely to be

correlated. This intra-cluster correlation must be taken into account in the calculation of sample size and in the analyses. We computed the number of patients per clinic, m, required for the study, using the following equation:[82]

$$m = \frac{1-\rho}{\frac{k}{n^*} - \rho}$$

where ρ = intra-practice correlation
k = number of clinics
mk = total sample size
n* = total effective sample size = mk/DE
and DE = design effect = 1 + ρ (m-1)

Using 6 clinics and assuming intra-class correlation of 0.014, we need to recruit 224 families with a child between 0 and 18 per clinic. We anticipate that drop-out is not a major problem and unlikely to be more than 10% for the primary outcome. We will account for this factor by oversampling patients by 10%. According to the data presented in Table 3 and with an assumption that for each category of FGID diagnosis 50% of the patients will be in the control clinics and other 50% will be in intervention clinics, there will be about 5,500 patients/year/group with a FGID diagnosis. If we assume that only 66% of the parents will respond to the phone call at 3-month post baseline, we will have 3,630 patients/year/group. Therefore, the number of patients from which we are able to recruit is not a factor at all. We can have more than 90% power even to detect a 5% absolute difference with all available patients in one year.

Data Analysis. Baseline clinical and demographic data will be compared between the intervention and the control groups. Categorical variables will be examined using chi-square tests and continuous measures with Student's t-tests. The statistical analysis of the data will be performed using the SAS.

Primary Outcome Analysis: The primary outcome is whether there was resolution of symptoms of a FGID by Rome IV criteria in children and adolescents or whether parental reassurance was achieved for infants. For this dichotomous outcome, we will use a non-linear mixed-effects model and will adjust for potential confounders (fixed effects) and clinic effects (random effects). To be specific, if Y_{ik} is a binary variable indicating whether the k^{th} patient in the i^{th} clinic was reported a resolution of FGID symptom, we have

$$\log it(Y_{ik} = 1) = \alpha_i + \mathbf{X}_{ik}\boldsymbol{\beta}, \text{where } \alpha_i \sim N(0, \varepsilon^2).$$

The random component α_i correlates outcomes of patients cared by the same clinic. Physicians would not be considered as a nesting factor since the same patients at different visits may not be seen by the same physician. The vector X_{ik} represents fixed covariates and β represents the corresponding coefficients. The fixed covariates will include critical covariates such as treatment group, gender, and other patient and clinic level characteristics. We will also examine potential interaction terms among the covariates, particularly interactions between treatment arm and other covariates. If there is a significant treatment difference, we will report the 95% confidence interval of the treatment difference after adjusting for all confounding factors. The model will be fitted by the SAS procedure NLMIXED.

Secondary Outcome Analyses: The additional time points of 1, 6, and 12 months will be modeled identically to the primary outcome analysis. Furthermore, we will have a number of secondary outcomes which are count data aggregated within the 12 months after initial Rome IV

diagnosis. We will use Poisson regression to model these count data. To adjust for the possible differences between the intervention and control groups, we will use an approach similar to the analyses of primary outcomes and again clinic will be considered as a cluster factor to take into consideration the correlation in count data within each clinic. Family and provider satisfaction scores will be analyzed using Student's t-test, as well as a linear mixed-effects model as with other outcomes.

Missing Data Analysis: We anticipate a low rate of missing data for the primary outcome since our CHICA system showed only 1.3% error rate due to omission of a data field. However, we anticipate a possible dropout rate of 10% for which we have adjusted the sample size. It is very unlikely that the dropout mechanism will depend on the unobserved outcome. Even if MNAR is true for certain subjects; given the low percentage of missing values, we think the bias induced is negligible. Under circumstances where compromise of power due to missing values is of concern, we will use a multiple imputation procedure to make use of all relevant observed data under the assumption that the data are Missing At Random (MAR). The SAS procedure MI and MIANALYZE will be used to implement the procedure.