Study Protocol and Statistical Analysis Plan (SAP)

MULTI-CENTER RANDOMIZED CONTROLLED TRIAL OF REFEEDING IN ANOREXIA NERVOSA:
The Study of Refeeding to Optimize iNpatient Gains (StRONG)

Principal Investigators	Andrea Garber, PhD, RD Professor of Pediatrics Division of Adolescent & Young Adult Medicine UCSF Benioff Children's Hospital University of California, San Francisco 3333 California St, Suite 245, Box 0503 San Francisco, CA 94143 ph: 415-514-2180 Neville Golden, MD Chief, Division of Adolescent Medicine Marron and Mary Elizabeth Kendrick Professor in Pediatrics Stanford University School of Medicine 770 Welch Road, Suite 100 Palo Alto, CA 94304 ph: 650-736-9557
Protocol identification number	UCSF IRB # 15-16368
ClinicalTrials.gov identifier	NCT02488109
Author	Andrea Garber
Version (Date)	1.3 (July 31, 2019)
Senior Statistician	Jing Cheng, MD, PhD Professor UCSF School of Dentistry Department of Preventive & Restorative Dental Science University of California, San Francisco 707 Parnassus Ave, 1026, Box 1361 San Francisco, CA 94143 ph: 415-502-0129

Abbreviations

AAN	Atypical anorexia nervosa
AN	Anorexia nervosa
DCC	Data Coordination Center
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, fifth edition
EDE	Eating Disorder Examination
EDE-Q	Eating Disorder Examination-Questionnaire
HCR	Higher Calorie Refeeding
kcal	Calories
LCR	Lower Calorie Refeeding
mBMI	median Body Mass Index (for age and sex)
SOC	Standard of Care

Table of Contents

PROTOCOL Version 1.3 (See Table of Amendments)

Multi-Center Randomized Controlled Trial of Refeeding in Anorexia Nervosa

1. Introduction

Current recommendations to guide the clinical care of AN patients hospitalized with medical instability due to malnutrition are based solely on retrospective or observational studies and/or clinical experience. No studies to date have prospectively tested high calorie refeeding (HCR) and the long-term impact on recovery is unknown. Consensus has developed over recent decades that patient safety can only be guaranteed using low calorie refeeding (LCR). The entrenchment of clinical practice without supporting evidence is a widely recognized dilemma in healthcare. While RCTs are considered the "gold standard" to establish evidence-based medicine, until recently there was insufficient data to propose such a study of refeeding in AN. We now have preliminary findings to indicate that LCR might be too cautious and that HCR appears feasible and may improve long-term recovery. Thus, we are poised to compare these two treatments in a parallel, randomized fashion.

2. Study Design

The purpose of this multi-center randomized controlled trial is to compare LCR vs. HCR refeeding strategies for hospitalized adolescents with AN. Participants will be recruited upon hospital admission at two centers (UCSF and Stanford) to maximize sample size, and randomly assigned 1:1 within site to one of the two strategies. A total of 120 participants age 12-24 yrs who meet DSM-5 diagnostic criteria for AN and atypical AN and present as medically unstable due to malnutrition will be enrolled. Treatments will not be blinded, since both the patients and clinicians who work with this population are highly skilled at estimating kcal and would be able to determine their group assignment by simply viewing the meal trays. In addition, target kcal will be reached faster in HCR and this would be apparent on physician orders. The proposed study is powered to detect a meaningful difference in clinical remission (Aim 1A).

2.1. Study Population

a. Inclusion/Exclusion: Adolescents hospitalized for medical instability secondary to malnutrition will be eligible as follows. Inclusion criteria: diagnosis of AN, atypical AN, age 12-24 years, no hospital admissions for the previous six months, and meet hospitalization criteria (daytime HR < 50 bpm or night time HR < 45 bpm, BP <90/45 mmHg, temperature < 35.6° C or orthostasis defined by increase in HR > 35 bpm or decrease in systolic BP > 20 mmHg from lying to standing). Exclusion criteria: diagnosis of bulimia nervosa [DSM-5], currently in remission (as defined by weight and EDE-Q score per Aim 1), admission for food refusal without malnutrition, current pregnancy, chronic disease (e.g. immune/endocrine disorders, pulmonary, cardiac, or renal disease), current suicidality or psychosis.

b. Participant recruitment and consent: Participants and their parents will sign assent (for those < 18 yr) and consent, respectively within 24 hr of hospitalization. This may occur in the clinic when they are deemed medically unstable and waiting transfer to the hospital or in the hospital if admitted directly. Consent will include permission to review all medical records, to review hospital billing data, and to contact for future research projects. If a participant turns 18 yr while enrolled, s/he will be reconsented. Participants 18 years of age and older are able to consent themselves; thus we will request verbal consent from their parents to complete parent surveys.

2.2. Randomization

Participants will be stratified by site and randomly assigned 1:1 to the two intervention strategies within 24 hr of hospital admission. The Data Coordination Center (DCC) will provide a secure unpredictable allocation sequence (e.g. A B B A …) which will be programmed into a secure electronic study tracking system for assignment of each accrued participant. The sequences will be generated using block size of two to four to maximize balance between arms throughout accrual while ensuring the sequences remain unpredictable. As patients consent to study participation, clinical-research staff will assign the next available study ID number in sequence, identify the allocated intervention arm, and inform participants and their families of the assignment. In turn, the clinical research staff will provide the linked study ID number and Medical Record Number to the DCC, which will store this Personal Health Information in a HIPAA-compliant manner along with the intervention assignment.

3. Study Procedures

3.1. Treatment and Follow-up

a. Study groups: Upon randomization, the Lower Calorie Refeeding (LCR) group will begin with 1400 kcal per day; Higher Calorie Refeeding (HCR) group will commence at 2000 kcal. Given our previous findings, recognition of the so-called underfeeding syndrome and recent clinical experience, we will not test a 1200 calorie diet even though it is still currently recommended. Our previous studies have adequately demonstrated that a 1200 kcal diet produces initial weight loss and therefore do not feel it is ethical to assign participants to this treatment. We chose 1400 kcal to start because in our previously study of LCR, weight loss ceased on day 3 in hospital when diets averaging 1411(299) kcal were prescribed. Diet prescriptions will increase by 200 kcal every other day in LCR and 200 kcal per day in HCR until a target level is reached. Target kcal are calculated upon admission as percent of energy needs using Estimated Energy Requirement (EER) equations from the Institute of Medicine. These equations are used clinically to set goals for caloric advancement although they are known to underestimate energy needs in patients with anorexia. Therefore, we maximize these estimations by using target weight corresponding to the mBMI for age and sex (rather than current weight), a moderate activity factor of 1.2-1.3 (despite bed rest), and additional 500 kcal (if current weight < the MBMI).

b. Intervention (refeeding protocols): During hospitalization, participants will follow a meal-based refeeding protocol that calls for eating three meals and three snacks per day, served on trays at the bedside, in the presence of 'Room Sitters'. The calorie level of the diet will be prescribed by the physicians per study protocol and the meals will be prepared by hospital foodservice. The study PI (Garber) will work with the Research Registered Dietitian (RD) and nutrition staff at both sites to ensure that menu selections fit the general macronutrient distribution of 30-40% fat, 15-25% protein and 35-55% carbohydrate. Menus will be continually analyzed (Software v.17.9.5, Computrition, Inc., Chatsworth, CA) to ensure conformance to this distribution as menu items are added or change over time. Dietetic Technicians will keep daily calorie counts, per Standard of Care (SOC), showing actual kcal consumed from food and formula.

   A high energy liquid supplement ("formula") providing 1.5 kcal per mL (360 kcal per 240 mL can) will be used orally as needed to replace kcal refused in meals or snacks per a standard calorie replacement protocol. We previously reported a greater than 98% concordance between kcal prescribed and actual kcal ingested using this method, as well as an equal proportion of kcal intake from formula in LCR and HCR groups. This finding supports our clinical observation that HCR meals can be completed without additional reliance on drinking formula. However, most AN patients do experience discomfort during refeeding and therefore all participants will receive SOC meal support including emotional support and techniques such as distraction. All beverages will be weighed and measured before placement on the tray, with a 1.5 L per day free water restriction. Room sitters will observe intake of all meals/snacks and remain in the room for 45 min afterwards.

c. Monitoring of electrolytes: Blood for electrolytes will be obtained between 5 and 7 am every 24hr for the first 7 days and more frequently if needed. Since the risk of refeeding decreases after the first week, electrolytes will be monitored every other day starting on day 8 unless there is continued evidence of abnormalities.

d. Correction of electrolyte abnormalities: Electrolyte abnormalities (serum phosphorus < 3 mg/dL, magnesium ≤ 1.7 mg/dL, or potassium < 3.5 mEq/L) will be corrected with a standardized protocol for both sites. Patients with hypophosphatemia will be treated with sodium potassium phosphate, 250 mg per packet (8 mmol phosphorus, 7.1 mEq potassium), one packet three times a day by mouth for a serum phosphorus between 2.5 and 2.9 mg/dL, two packets (500 mg) three times a day for a serum phosphorus of 2.0 - 2.5 mg/dL, and if serum phosphorus is < 2.0 mg/dL then intravenous sodium potassium phosphorus will be initiated at a dose of 0.24 mmol/kg (max of 15 mmol per dose). The PICU will be contacted and labs will be rechecked STAT 4 hours after the infusion is completed. Those with hypomagnesemia will be prescribed magnesium oxide (150 mg elemental Mg per tablet), one tablet three times a day by mouth for a serum magnesium between 1.3-1.7 mg/dL; two tablets three times a day by mouth for a serum magnesium between 1.0-1.2 mg/dL. In the case of serum magnesium below 1.0 mg/dL, the PICU will be called for possible transfer, and intravenous magnesium sulfate will be started at a dose of 50 mg/kg (max of 2 grams per dose). Labs will be rechecked STAT 2 hours after infusion completed: If repeat serum mag still < 1.0 mg/dL, repeat same dose of magnesium sulfate IV; If repeat serum mag > 1.0, begin PO Mag, at PO dose indicated above. Those with hypokalemia will be prescribed extended release potassium chloride by mouth, 20 mEq for a serum potassium of 3.1-3.4 mmol/L; and 40 mEq for a serum potassium of 2.5-3.0 mmol/L. For a serum potassium between 2.2 - 2.5 mmol/L, 40 mEq extended release potassium chloride will be given STAT and the PICU will be called for possible transfer. Labs will be rechecked in 4 hours (peak) and 12 hours (estimated nadir). For any serum value of potassium < 2.2 mmol/L, intravenous potassium chloride will be initated and PICU will be called for transfer. Declining electrolyte levels that are in the normal range will not be treated.
Participants will also receive a SOC supplement regimen including 500mg elemental calcium with vitamin D twice per day and an adult multivitamin with minerals once per day.

e. Study time points: Participants will be followed prospectively in hospital with daily measures of calorie and supplement intake and weight from admission through discharge. Patients will be discharged when medically stable, with the primary criterion of heart rate ≥ 45 bpm for least 24 hr. Full medical stability defined as HR ≥ 45 bpm for 24 hrs, temperature ≥ 35.6ºC for 24 hrs, ≥ 75% of mBMI, BP ≥ 90/45 mmHg for 24 hrs or if systolic BP < 90 then asymptomatic and all else stable, orthostatic change in HR ≤ 35 bpm or if > 35 then asymptomatic and all else stable; and orthostatic change in SBP ≤ 20 mmHg or if > 20 then asymptomatic and all else stable.

Timing of Procedures: Table 1

	INPATIENT			OUTPATIENT FOLLOW---UP				
	Admit	Daily	DC*	10 dy	1 Mo	3 Mo	6 Mo	12 Mo
MEDICINE/NURSING PROCEDURES								
Weight	SOC	SOC	SOC	X	X	X	X	X
Height	SOC			X	X	X	X	X
Vital Signs	SOC	SOC	SOC	X	X	X	X	X
Electrolyte monitoring ¥	SOC	X	SOC					
NUTRITION PROCEDURES								
24---hr food recall	X			X	X	X	X	X
QUESTIONNAIRE ADMINISTRATION								
EDE-Q	X		X		X	X	X	X
HCUMS survey				X	X	X	X	X
Demogr & Eating Disorder	X							

Follow---up form				X	X	X	X	X

* DC = discharge; SOC = Standard Of Care; Rnd = Randomization; ¥ SOC is every other day; we will monitor daily as part of Aim 2.

3.2. Data Collection

Other than the treatment we are testing (HCR vs. LCR), patients will receive SOC in the hospital. Thus, as shown in **Table 1**, the vast majority of procedures in hospital are SOC. Follow-up visits, on the other hand, are for the purpose of collecting data and will be scheduled at the designated time points.

a. **Baseline data collection:** The following covariates will be collected upon admission, prior to randomization.
   (1) Demographics and eating disorder history: an intake form will be self-administered (with study coordinator as proctor, 15 min) to assess: highest body weight, lowest body, date of onset (to calculate length of illness and rapidity of weight loss), family history of eating disorder, self-reported race/ethnicity, maternal education and zipcode (to indicate socioeconomic status), date of birth.
   (2) Eating Disorder Examination-Questionnaire (EDE-Q): is a standardized research interview that measures eating disorders psychopathology. Dr. Le Grange (co-I, UCSF) has used this tool extensively in RCTs examining psychotherapeutic modalities and long-term recovery in AN and BN and to categorize lower and higher risk study participants. He will oversee the psychological aspects of this study, including the EDE-Q.
   (3) Food recall: dietary intake for the day prior to hospital admission will be assessed with a 24-hr food recall by the Research RD and analyzed via Nutrition Data System for Research (NDS-R) for total kcal and macronutrient profile. Dietary Density (DDS) and Variety Scores (DVS) will be calculated since acceptance of more energy dense and variety of foods in hospital has been shown to predict recovery at 8 mo.
   (4) Severity of illness: %mBMI and HR on admission
   (5) Health Care Utilization and Missed School (HCUMS) Survey: This proctored interview draws from established tools to assess cost, and has been tailored for AN and integrated with the follow-up form (above). In addition to to "other care", it assesses health care utilization in the 6 months prior to admission including medications, physician visits, dental visits, ER visits and laboratory testing. It will be administered by trained research staff, who will document parent- and participant-reported care, supplemented with dates, doses, and other details available in the medical record.

b. **Data collection daily in Hospital:**

   (1) SOC in hospital: Participants will be followed daily in hospital. Consistent with SOC for these patients, night time HR will be assessed with continuous cardiac monitoring throughout hospital stay, temperature will be measured orally and BP will be measured every 4-8 hr. Postural changes will be assessed with supine measurements (after 5 min), followed by standing measurements (after 2 min). When multiple vital signs measures are taken per protocol during one hospital day or one outpatient clinic visit, the most deviant value (lowest HR, lowest BP, greatest increase in HR and greatest decrease in BP on orthostatic changes, lowest Temp) will be recorded. Weight is measured every morning upon waking after voiding on an electronic scale, with the subject wearing only a hospital gown. Height will be measured within 24 hours of admission with wall-mounted stadiometer. Electrolytes will be monitored per SOC as described above in 3.1.c.

   (2) Study in hospital: Participants at both sites will complete the questionnaires above.

c. **Data collection during follow-up:** Participants will return for five study follow-up visits through 12

months post-discharge and complete the following procedures:

(1) Anthropometrics and vital signs: Height, weight and vital signs will be measured according to the in-hospital protocol by trained medical staff with standard equipment. Vital sign measures will be taken after a 20-min rest to minimize the influence of activity required to attend the visit (e.g. walk from car). After resting, vital signs will be measured in the research center with standard, calibrated equipment and with postural changes according to the procedure above. Data will be entered directly into the electronic data capture system with fillable and constrained sections for anthropometric measures and vital signs to minimize error.
(2) Food recall: dietary intake for one full day during the week of each follow-up visit will be assessed with a 24-hr food recall by the Research Dietitian.
(3) HCUMS Survey: As described above, this proctored survey will assess utilization of health care such as re-hospitalizations since the time of the study-related hospitalization, participation in eating disorder treatment programs, medications, missed school, missed work, and other direct and indirect costs associated with eating disorder care since the time of each prior follow-up visit. The HCUMS will document medications pertinent to recovery measures (menses and psychopathology), including medications, current mental health care, other medical and psychological/psychiatric care ("other care") outside of our medical centers (e.g. residential care, psychiatric hospitalization). Psychotherapy modality and adherence may be important prognostic covariates of long-term outcomes in this open follow-up study.

3.3. Safety

This study begins with a hospitalization as per SOC for patients who are medically unstable with malnutrition secondary to AN. Patients will be admitted to the adolescent medicine service if they are deemed medically unstable per published criteria. Once admitted, patients will be eligible for study enrollment. The treatment (HCR or LCR) is limited to the hospital stay. Aside from the questionnaires at both sites and daily (instead of every other day) electrolyte monitoring, all hospital procedures are consistent with SOC. After discharge, participants will be followed openly. They are required to be under a physician's care to ensure medical stability but not required to receive that care from us (however many do). Many patients have a psychiatrist to manage psychiatric co-morbidities such as anxiety and depression. If they receive care or hospitalization elsewhere they can still continue in the study and we will collect that with our follow-up form.

a. Prospective monitoring of AEs: Aim 2 specifies three electrolyte abnormalities that will be monitored prospectively in all participants and documented as described in **3.1.c.&d.**: hypophosphatemia (<3 mg/dL), 2B) hypomagnesaemia (≤1.7 mg/dL), and 2C) hypokalemia (<3.5 mEq/L).

b. Data and Safety Monitoring Board (DSMB): As a multi-center clinical trial comparing treatments, the proposed study is required to have a DSMB according to the NICHD policy for clinical research monitoring. The purpose of the DSMB is to ensure the safety of participants and validity of the trial. We will draft a DSM Plan using the NICHD template.

STATISTICAL ANALYSIS PLAN (most recent change May 26, 2019)

Multi-Center Randomized Controlled Trial of Refeeding in Anorexia Nervosa

1. Aims and Objectives

Our study has three main aims. We will compare:

AIM 1: Efficacy of LCR vs. HCR. We hypothesize that LCR and HCR will differ by *achievement and maintenance of*: (1A) clinical remission during 12 mo follow-up, defined by achievement of both (i) weight ≥ 95% median BMI (MBMI) for age and sex, and (ii) Eating Disorder Examination-Questionnaire (EDE-Q) global score within 1 SD of clinical norm, and (1B) medical stability during initial hospitalization, defined by published vital sign thresholds.

AIM 2: Safety of LCR vs. HCR during initial hospitalization. We hypothesize that LCR and HCR will *not* differ by *incidence of*: 2A) hypophosphatemia (<3 mg/dL), 2B) hypomagnesaemia (≤1.7 mg/dL), and 2C) hypokalemia (<3.5 mEq/L).

AIM 3: Cost-effectiveness (CE) of LCR vs. HCR. We hypothesize that HCR will be more cost-effective than LCR, as determined by cost (including costs of initial and re-hospitalizations, 12 mo follow-up, other care, and safety/adverse events (AEs) and effectiveness per adolescent recovered (defined in AIM 1A and B).

2. Statistical Methods

2.1. Pool of participants

a. Projected pool of eligible participants and accrual rate: At Stanford, in 2012 there were 295 admissions to the dedicated inpatient eating disorders unit, with approximately 36% of patients meeting DSM-4 criteria for AN; 100 similar patients were admitted at UCSF. With the broader eligibility criteria also including DSM-5, we anticipate at least 40% of patients (120 per yr at Stanford and 40 at UCSF) will be eligible. Of those who are eligible, we estimate that at least 50% will agree to participate and thus we will not attempt to achieve equal enrollment across sites. We will accrue 3-4 participants per mo over 3 yr until N=120 is reached and retain 85% of this sample through 12 mo as shown in open follow-up studies of participants with AN.

b. Commitments of Site PIs, Research Teams, and Participants: Both site PIs have successfully recruited and retained AN participants in research projects and seen them to completion and publication. Furthermore, co-I Cheng is a faculty biostatistician with extensive experience in clinical trials. She will lead the DCC, aiming to ensure the trial is designed, executed, and analyzed without bias. Patient incentives to participate will emphasize the value of their contributions to medical research and remuneration for their time.

c. Data Analyses:

Sample Description: The study sample will be summarized and described (e.g., mean ± SD) by stratification factor and baseline covariates to confirm general balance by arm and data will be summarized for completeness of follow-up (e.g., length of stay, last visit).

2.2. Outcomes

The primary analysis of the study adopts a modified intent-to-treat (mITT) approach to compare outcomes between randomized HCR and LCR participants who received treatment for at least one day. Patients who are ineligible post randomization, provide no assent after parent's consent, or withdraw before receiving any treatment, will not be included in the mITT analysis for reasons of data unavailability, ethics and clinical relevance. Sensitivity analyses will be conducted to check that 1) withdrawal patients are not different from patients in both groups included in mITT in baseline covariates; and 2) including withdrawal patients in the analyses will not change results and conclusion of mITT analyses. The sensitivity analyses will provide us

reassurance of mITT results.

Aim 1A: Primary (long-term) efficacy outcome. A (generalized) linear mixed-effects regression model (GLMM) will compare study arms with respect to achievement and maintenance of clinical remission. Clinical remission is defined as the combination of mBMI and EDE-Q score at mo 1,3,6,12 (separate analyses below) and measured at 1, 3, 6, and 12 months as 1) clinically remitted (yes or no) and 2) weight recovered (percent of mBMI). The models will include time, treatment group, time*treatment group interaction, and unbalanced baseline covariates if any as fixed effects, while sites and patients will be included as random effects to account for the correlation due to clustering. The time*treatment group interaction provides mITT effect of HCR compared to LCR on clinical remission over time. The average remission rates and scores and their 95% confidence intervals (CIs) will be estimated from the model. GLMM accounts for the fluctuating nature of mBMI and EDE-Q in AN and uses all available data with missing at random assumption, instead of a stronger assumption of missing completely at random required in other models. Aim 1 models will be supplemented with secondary analyses: (i) potential moderators at baseline (included in regression models as main effects and interactions with period and time), may include DSM-5 criteria and EDE-Q thresholds of risk and (ii) potential mediators at follow-up (included as time-dependent covariates), may include food recall (DDS and DVS), healthcare utilization, or incidence of AEs. In addition, separate mixed-effects models will analyze continuous versions of mBMI and EDE-Q to describe longitudinal trajectories.

Aim 1B: Secondary (short-term) efficacy outcome: Time to restore and maintain medical stability in hospital is defined as days to reverse the medical instability indicators for hospitalization in adolescents with eating disorders. A six-point index will adjudicate daily medical stability: 1.) 24-hour heart rate (HR) ≥ 45 bpm, 2.) systolic blood pressure (SBP) ≥ 90 mmHg, 3.) temperature ≥ 35.6° C, 4.) orthostatic increase in HR ≤ 35 bpm, 5.) orthostatic decrease in systolic BP ≤ 20 mmHg, and 6.) ≥ 75% of mBMI. Each of the six criteria were scored as "1" if met, "0" if unmet and missing (not scored) if not measured. Medical stability was considered restored when all criteria were stable for 24 hours, allowing a maximum of two missing values (i.e. participants were considered stable if meeting 4 of 4, 5 of 5 or 6 of 6 measured criteria). All randomized participants who receive at least one day of treatment, including those who withdraw at any time during the refeeding intervention, will be included in the mITT analysis. Specifically, survival analysis with log rank test will compare time to achieve medical stability by arm while accounting for the correlation within sites; participants who do not meet stability criteria by hospital discharge will be right-censored. In case of any important unbalanced covariate at baseline, Cox proportional hazard ratio model will be used to control for the potential bias due to the confounder. Additionally efficacy outcomes will include the proportion achieving medical stability in each arm, change in %mBMI as compared to baseline, and time to restore HR.

Aim 2. Safety: Primary indicator of safety will be incidence of the following electrolyte abnormalities during hospitalization: 2A) hypophosphatemia (<3 mg/dL), 2B) hypomagnesaemia (≤1.7 mg/dL), 2C) hypokalemia (< 3.5 mEq/L). These AEs will be tracked, recorded, reported to the DSMB for monitoring, and to the IRB and NIH as needed, according to the attached DSMP. Incidence of the electrolyte abnormalities and proportion receiving supplementation to correct electrolyte abnormalities will be compared by arm during hospitalization using Fisher's exact test. Additional safety outcomes: the proportion receiving electrolyte supplementation during the hospitalization between groups will be compared with Fisher's exact test; time to electrolyte nadir will be compared with Wilcoxon rank sum test.

Aim 3. Cost Effectiveness (CE): A decision tree of treatment costs, AEs, health care utilized (including rehospitalizations), and remission will compare the CE between the two study arms. The main CE outcome is incremental cost per additional adolescent remitted. Health care utilization will be costed using national data sources such as acquisition costs for medications, Kids-HCUP for hospitalizations, CPT codes for physician visits and mental health visits, and internet-based costs for lab tests. We will use 2014 US costs and not charges. Indirect costs including missed school and workdays will be assessed and costed using national estimates of wages and salaries of this age group from the Bureau of Labor Statistics, in order to estimate loss of leisure time (school) or salary

time (work).
We will determine the incremental CE ratio (ICER) as:

ICER = CostLCR-CostHCR/Number RecoveredLCR-Number RecoveredHCR.

Effectiveness will also be indicated by cost of rehospitalizations avoided (ie, rehospitalizations in LCR-HCR). We will determine the net monetary benefit (NB) of each treatment option as: NB=Effectiveness X Willingness To Pay (WTP) – Cost. A positive difference in NB between treatments indicates CE. We will also calculate an acceptability curve to demonstrate how parameter uncertainty affects the likelihood of selecting the optimal treatment at a given WTP threshold. Cost of treatment will be determined by initial hospitalization and 12 mo follow-up costs (not charges) including AEs and rehospitalizations. The HCUMS follow-up survey will assess indirect costs such as lost school and/or work (wages) using national data sources (**see C.2.c.(2))**.

Effectiveness will be determined per aim 1A; ICER will also use time (incremental cost per additional day of recovery time over 12 mo). Cost efficacy will be assessed with short-term outcomes at end of hospitalization. Per Aim 1B, cost-efficacy will be assessed with other short-term outcomes at the end of treatment (hospitalization period). Cost-efficacy will be defined as hospital cost or charges associated with length of stay; group differences will be compared with Wilcoxon rank sum test.

2.3. Power and Sample-size Considerations

Aim 1. Based on studies of AN remission , **Table 2.A** shows that with N=60 per arm we have 80% power on 2-sided 0.05-level test to detect a 20% difference (8% vs. 28%) if data were cross-sectional (ρ=1). Our longitudinal data will allow detection of smaller effects, especially if the correlation among outcomes is low (ρ <0.8). We anticipate 85% retention and non-differential dropout by arm. Since **time to medical stability** is also expected to differ by at least 12% (**Table 2.B**.), we will be adequately powered for Aim 1B.

Table 2.A Detectable differences in remission---rates:

ρ*	LCR	HCR	Mo---3 Diff
1.0	8%	28%	20%
0.8	14%	34%	20%
0.6	16%	34%	18%
0.4	18%	34%	16%
0.2	20%	34%	14%
0.1	22%	34%	12%

Table 2.B. Detectable differences in time to medical stability-rates

ρ*	LCR	HCR	Difference
1.0	72%	92%	20%
0.8	66%	86%	20%
0.6	66%	84%	18%
0.4	66%	82%	16%
0.2	66%	80%	14%
0.1	66%	78%	12%

*correlation among 5 time points within participant

2.4. Data Management

a. DCC: Dr. Cheng at UCSF will lead the DCC, which will be autonomous and independent of the clinical sites. It is housed in the UCSF Department of Preventive and Restorative Dental Sciences (PRDS). The department is staffed primarily by statisticians and scientific researchers who conduct data-intensive research, collaborative data collection and analyses from multiple sites. They are equipped with an independent network of sophisticated and reliable computer systems with high-level security for protecting health information. The network is maintained by an in-house computer staff, which manages all aspects of the network, including ongoing maintenance, installation and upgrades of hardware, software and structural components such as cabling and servers. Dr. Cheng is the lead biostatistician and routinely guides the work of Master-level statisticians, data managers and programmers. **Dr. Cheng** will continue as faculty biostatistician and DCC leader. She has extensive experience running NIDCR-funded DCCs for clinical trials with more than 8-10 sites nationwide. Thus, UCSF has experience maintaining a distinctly separate but closely coordinated working relationship between clinical sites and data center.

b. Electronic data capture: Both clinical sites are equipped with the same Research Electronic Data

Capture (Qualtrics) system for databases, data entry forms, online questionnaires and data validation. Data will be automatically exported to STATA or SAS for analysis using The Data Export Utility. The DCC uses advanced features including branching logic for dynamic data entry form generation, file uploading, data importing, and embedded calculated database fields.

c. Confidentiality: Loss of confidentiality is a recognized risk of participating in clinical research since protected health information, medical history, and demographics are used for the study. Loss of privacy may lead to problems with insurability or social stigmatization. We will make effort to minimize this risk and have systems in place to ensure confidentiality. Data will be de-identified and thereafter handled by ID number, rather than by name. No publications will include the names of patients or identifying information about study participants.

2.5. Retention and Attrition

We expect to retain 85% of our sample through one year of follow-up. This is consistent with other open follow-up studies of AN and feasible given our patients volumes and return rates. We will actively retain participants by providing incentives: movie tickets upon enrollment and a $50 for every follow-up visit attended. Primary analyses will use intent-to-treat longitudinal models that will include outcomes from randomization through the time of dropout or 12 mo, whichever is longer. Secondary analyses will adjust models for baseline covariates that may be associated with loss to follow-up. We anticipate very few missing outcomes because weight and vital sign (medical stability) measures are SOC in AN care during hospital and at follow-up and the majority of patients hospitalized at our programs return to us for follow-up care. Reasons for refusal to participate will be collected from patients and families who decline enrollment.

APPENDIX

MULTI-CENTER RANDOMIZED CONTROLLED TRIAL OF REFEEDING IN ANOREXIA NERVOSA:
The Study of Refeeding to Optimize iNpatient Gains (StRONG)

Appendix Table 1. Amendments to the Protocol

Protocol version	Section changed	Amendment description & reason	IRB approval and/or other documentation date
1.1	3.1 Treatment and Follow-up	Oral nutrition supplement to replace refused food contains 1.5 kcal/mL; previously listed as 1.06 kcal/mL, to reflect actual clinical practice	IRB approved 12/01/2015
	3.1 Treatment and Follow-up	Electrolyte replacement protocol developed for more severe levels of low serum electrolytes	
	3.1 Treatment and follow-up	Nutrition intake is assessed via 24-hour recalls at each follow-up time point to reduce participant burden	
	2.1 Study Population	Added one exclusion criterion: extremely malnourished patients, admitted with mBMI <60%, will no longer be eligible to enroll in the study, as they are at higher risk of medical decompensation.	Posted on clinicaltrials.gov 05/12/2016
	3.1 Treatment and follow-up	Follow-up visits will be defined as time since discharge instead of time since admission. This will ensure that time elapsed will be comparable among participants despite differing lengths of admission.	
1.2	3.2 Data Collection	Self-reported EDE-Q (5-10 minutes) will replace the lengthy interview-based EDE (60-90 minutes), based on recent evidence that it serves as a valid proxy for measuring severity of eating disorder psychopathology, with significant reduction in participant burden and personnel cost.	IRB approved 01/05/2016 NIH in progress report 05/08/16 Posted on Clinicaltrials.gov 05/12/2016
	2.1 Study Population	Eligibility criteria updated to reflect actual clinical practice: to be eligible for participation, orthostatic increase HR from lying to standing of 35 bpm	IRB approved 05/03/2016
1.3	3.1 Treatment and follow-up	Discharge criteria based on SBP only, not DBP	IRB approved 05/19/2017

Appendix Table 2. Amendments to the Statistical Analysis Plan*
*** All amendments were made and documented prior to locking of database for analysis**

Section changed	Amendment description & rationale	Amendment date & documentation
2.2 Outcomes	**Amendment 1: End-point for AIM 1B shortened to in hospital treatment period** **Rationale:** To match the timeframe for the safety outcome and to better capture the efficacy of this relatively short intervention (less than 2 weeks in hospital) within a long, open follow-up trial. Further, to allow timely dissemination of findings. [Citation: IOM guideline, *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk,* pg 118: Investigators "may publish the primary trial endpoints despite ongoing longer-term participant follow-up; in this case, the last participant's last visit may not occur for some time, and hence the full analyzable data set may not be complete at the time of the original publication."] Timeframe for AIM 1B specified as 2 weeks and posted on clinicaltrials.gov Project officer, Dr. Karen Winer approved analysis of short-term (in-hospital) outcomes by arm (DSMB notified 7/03/19, randomization code was broken) Short-term database locked	 06/30/2015 05/26/2019 07/31/2019
2.2 Outcomes	**Amendment 2: Analytic approach for AIM 1B changed from mixed effects regression modeling to survival analysis** **Rationale:** survival analysis chosen to examine time to restore medical stability in hospital (amended AIM 1B), rather than mixed effects regression modeling to examine medical stability over 12-month follow-up. This is a more appropriate approach to assess "time to" outcome, which accounts for the variable lengths of hospital stay and allows participants who were discharged before medical stability was restored to be right-censored. Change in "time to" analysis documented in NIH Progress Report Data Coordination Center documented decision to use survival analyses for AIM 1B	 05/05/2017 10/17/2017
2.2 Outcomes	**Amendment 3: Analytic approach for AIM 2 (safety) changed from Cox regression modeling to basic non-parametric testing** **Rationale:** Originally planned approach (Cox regression models) assumed high incidence electrolyte abnormalities based on prior studies reporting rates up to 48%. However, ongoing AE monitoring during the trial revealed very low incidence and therefore basic parametric methods were chosen to allow clinical interpretation of the results. DSMB reports since 2017 (signed by DSMB members and sent to POs, Dr. Graves and Dr. Winer)	 02/17/2017 02/20/2018 02/15/2019