R21 Environmental Enrichment Study Protocol

Translation of the EE intervention: The environmental enrichment (EE) paradigm was translated by a multidisciplinary team composed of investigators with expertise in psychology, physiology, neuroscience, integrative mind-body disciplines, gynecology, and translational research, together with consultants in HPA axis research and stress management, met during a period of one year to design an intervention to be tested in patients. A pilot randomized clinical trial was designed to test the efficacy of reducing painful symptoms and improving QoL as an adjuvant to standard care. The adaptation followed the ORBIT Model proposed by Czajkowski *et al.* (2015) to develop behavioral interventions for chronic diseases [refs]. This model guides the early pre-efficacy stages of the development of a behavioral intervention using a four-phase approach (F1: design, F2: preliminary efficacy, F3: efficacy, F4: effectiveness (F2 to be tested in the current proposal). Guided by the model, we translated a basic neurobehavioral finding (*stress control and EE reduces the growth of lesions in rats*) into a clinical question (*will an EE intervention be effective in reducing pain symptoms/inflammation and increasing QoL in women with endometriosis?).* First, we conducted a systematic literature review to select the approaches for the adapted EE intervention, using the PRISMA-P 2015 protocol and the PRISMA evaluation checklist. The process of translating the environmental enrichment (EE) paradigm has been described in detail in Nieves-Vázquez et al. (2022) [41].

Recruitment: After IRB approval (Protocol #1901004205R003) a recruitment campaign was conducted using social media (Facebook, Instagram, Twitter) of the Fundación Puertorriqueña de Pacientes con Endometriosis (ENDOPR), the only patient support association in Puerto Rico established in 2015. Recruitment for session one took place from September 2019 to December 2021 and then from May to June 2021. Recruitment for session two took place from August 2021 to February 2022. Patients interested in participating will be first screened for inclusion/exclusion criteria by our research coordinator and verbally consented to before being randomized into the intervention and control groups described below.

Inclusion Criteria: Participants were women with a surgical diagnosis of endometriosis, 18-50 years old, who are symptomatic and refractory to hormonal treatment, able to provide written informed consent, and can commit the time to participate in the environmental enrichment (EE) intervention for a period of 3 months. *Exclusion Criteria:* We excluded patients who present endometriosis symptoms but who have not been diagnosed by surgery by an OB-GYN specialist; pregnant women or those who plan to or become pregnant during the study period; women with endometriosis under hormonal treatment who report no symptoms as defined predominantly by pain; post-menopausal women; documented or visual cognitive or physical impairment that would interfere with participation or consent; currently under mental health treatment or using

steroid medications; affected by other confounding conditions, including pain syndromes (e.g., fibromyalgia, chronic fatigue syndrome, arthritis).

Study design of the pilot RCT: To evaluate the efficacy of a translated EE intervention for endometriosis patients, we conducted a RCT of parallel design with an intervention group (EE intervention) (n=29) and a wait list control group (n=27) from August 2021 to July 2022. Participants randomized to the intervention participated in six EE modules on alternate Saturday mornings. They could receive (or continue receiving) standard gynecological care (hormonal, analgesics, or surgeries) and psychological therapy as needed. Participants randomized to the wait-list control condition were invited to participate in an online seminar about endometriosis and could also receive standard of care for endometriosis and mental health. After providing informed consent, all subjects completed a set of validated surveys at baseline and end of the intervention to assess QoL, perceived stress, anxiety, depression, and pain symptoms, and provided saliva samples. Individuals in the intervention group were followed up 3-months after the study ended to assess possible long-term effects of the intervention. Participants in the control group completed the same questionnaires and provided the saliva samples during house visits that took place within one week before and one week after the timeline of the intervention. Clinical and socio-demographic data, as well as pain catastrophizing scores were obtained with the Endometriosis Phenome Project (EPHect) questionnaire [42]. Participants in the intervention group used WhatsApp chats to continue their communications between meetings. Data regarding treatments and doctor's visits during the study were obtained from all participants with a clinical history questionnaire administered twice during the study period.

Study surveys: Participants in the EE condition completed the following surveys at baseline and end of the intervention, and at 3 months after the intervention was completed. Individuals in the control group completed the same questionnaires only at baseline and end of the study period:

(1) *Endometriosis Phenome Project Clinical Minimal Questionnaire (EPQ-M):* The validated Spanish version of the EPhect's EPQ-M survey was used to collect cross-sectional, self-reported data. Developed by the World Endometriosis Research Foundation (WERF), this survey standardizes data collection from patients, including demographics, medical history, ob-gyn history, and lifestyle [42]. Different types of pelvic pain (dysmenorrhea, dyspareunia, chronic pelvic pain) are measured using a numerical rating scale from 1 (no pain) to 10 (worse pain) [43].

(2) *Endometriosis Health Profile-30 (EHP-30)*: This survey measures the endometriosis-related health status through 30 items covering five disease-related scales (core questionnaire): pain, control and helplessness, emotional wellbeing, social support, and self-image. EHP-30 has been shown to be sensitive to changes in patient outcomes, making it a useful tool in endometriosis clinical trials. Response categories are rated on a five-point Likert scale (0 to 4). The global QoL score is converted on a scale of 0 to 100, with the lower score representing a better quality of life [44, 45].

(3) Brief Pain Inventory (BPI): This questionnaire measures both pain intensity (minimum , maximum, average, current) using a numerical rating scale (NRS) of 0 to 10, with the highest number representing the worst imaginable pain. This instrument also measures the degree of pain impairment on function regarding general activity, mood, walking, work (including housework and paid work), relationships, sleep, and enjoyment of life. A global pain impact score can be calculated, with the higher score indicating worse impact [46].

(4) *Perceived Stress Scale (PSS-14):* This self-assessment tool measures the level of perceived stress due to life situations over the past month. It assesses to what extent respondents feel in control of unpredictable or unexpected situations, or conversely, whether they feel out of control and experience stress that leads to discomfort. It consists of 14 items with a response format of a five-point scale (0=never, 1=almost never, 2=occasionally, 3=often, 4=very often). A high total score corresponds to a high level of perceived stress [47].

(5) *Anxiety (General Anxiety Disorder-7 (GAD-7):* This self-administered questionnaire measures anxiety symptomatology through seven items which are scored from 0 to 3. The cut points of 5, 10, and 15 represent mild, moderate, and severe, respectively [48]. Patients reporting moderate-severe scores in GAD-7 were provided with a referral for psychological services.

(6) *Patient Health Questionnaire-8 (PHQ-8)*: This self-administered questionnaire measures depression symptomatology through eight items which are scored from zero to three. Scores of 5, 10, 15, and 20 represent cut points for mild, moderate, moderately severe, and severe, respectively. [49]. Patients reporting moderate-severe scores in PHQ-8 were provided with a referral for psychological services.

(7) *Pain catastrophizing scale (PCS)*: This scale is part of the EPhect-Q Questionnaire that was completed at baseline. The PCS evaluates three dimensions of catastrophizing: helplessness, rumination, and magnification [50]. The scale has a total score of 52 with items scored from 0 (‘not at all’) to 4 (‘all the time’) [51]. A PCS score higher than 30 is considered clinically significant and identifies those with a higher risk of chronicity and disability due to pain [52].

Saliva sampling and cortisol ELISA: Saliva samples for cortisol were collected at baseline, end of the intervention, and at the 3-month follow up for the experimental group. Saliva samples (approximately 1-3 mL) were obtained by passive drool into pre-labeled 15 mL tubes. These samples were obtained at the same time of the day (between 8:00-9:00am and at noon, before and after the intervention) to account for circadian variations in cortisol levels. For the control group, samples were collected one week before the start of the intervention and one week after during house visits. Control individuals were asked to collect saliva at the same time as the intervention group, on the same day, and store them at 4°C until collection during visits by the research team. Saliva samples were stored short-term on wet ice during transport to the research laboratory where they were processed and stored at -80°C until analysis. Once in the lab, the saliva samples

were spun down to clear buccal cells, and aliquots stored at -80°C until analysis by ELISA. Cortisol analysis was done using the High Sensitivity Salivary Cortisol Enzyme Immunoassay kit (Salimetrics, State College, PA, cat #130025), validated for the quantitative measurement of salivary cortisol, following the manufacturer's protocol. All samples were assayed undiluted in duplicate on the same assay plate. Cortisol concentrations were calculated based on standard curve, averaged, and reported in mg/dL.

Statistical analysis: The pilot trial's acceptability and feasibility were determined by rates of Recruitment, Enrollment, Adherence, Survey completion and Module evaluations.Equivalence of baseline characteristics, as well as between-group differences in clinical-demographics variables and study outcomes were assessed using descriptive and univariate statistics, including T-test or Mann-Whitney test for continuous variables (depending on normality of distribution, as assessed by Shapiro-Wilk test), and Chi-square or Fisher's exact test for categorical variables. Between group differences, as well as changes in the intervention group (from baseline) were evaluated for % improvement. Outcomes were evaluated using repeated two-way analysis of variance and assessed differences in intragroup (baseline vs. end of intervention) and intergroup (intervention vs. control). If findings were significant, we used a Tukey post-hoc test. Clinically meaningful changes in pain were considered when there was a change of three or more points (substantial difference), two to three points in the NRS (moderate difference) or 1 point (minimal difference) [43, 53]. A p value of 0.05 or less was considered statistically significant.