The American Journal of Human Genetics, Volume *110*

Supplemental information

The NYCKidSeq randomized controlled trial: Impact of GUÍA digitally enhanced genetic results disclosure in diverse families

Sabrina A. Suckiel, Nicole R. Kelly, Jacqueline A. Odgis, Katie M. Gallagher, Monisha Sebastin, Katherine E. Bonini, Priya N. Marathe, Kaitlyn Brown, Miranda Di Biase, Michelle A. Ramos, Jessica E. Rodriguez, Laura Scarimbolo, Beverly J. Insel, Kathleen D.M. Ferar, Randi E. Zinberg, George A. Diaz, John M. Greally, Noura S. Abul-Husn, Laurie J. Bauman, Bruce D. Gelb, Carol R. Horowitz, Melissa P. Wasserstein, and Eimear E. Kenny

Supplemental Methods

Data Extraction and Cleaning

Study data was collected and managed using REDCap (Research Electronic Data Capture), a secure web-based software platform hosted at the Icahn School of Medicine at Mount Sinai. Data collection instruments captured study activities and participant engagement, including but not limited to referral eligibility, recruitment, randomization, enrollment, consent and data sharing preferences, sample acquisition, study procedures, genetic findings, and survey assessments collected by clinical research coordinators and genetic counselors.

Data cleaning was conducted systematically throughout the project with final cleaning and data processing following the completion of study-specific time points, and at the end of data collection in May 2022. Data was queried for potential entry errors, missingness, outliers, and other quality assurance (QA) and quality control (QC) inquisitions ensuring the integrity of the data. Study staff reviewed queries and checked source material including REDCap instruments and audit logs, chart notes in the electronic medical record at pre-test and post-test, and paper versions of the phenotype checklists, testing requisition forms, genetic reports, and participant surveys. If the REDCap entry and source material did not match, the source material was used to reconcile the data. If the REDCap entry was missing, the source material was checked to determine if the value was an entry error or if it was a true missing. If the value was an input error, and verified by source, changes were made to the raw data in REDCap. If the value was truly missing, the entry would remain as such. If the REDCap entry matched the source material but was a clear error based on QA across another variable, decisions were made on how to handle these discrepancies. All queries, final resolutions and those that remained unresolved, and changes to the data were recorded in audit and change logs. Upon resolution of all data queries, raw data was exported from REDCap for analysis.

Participant Inclusions and Exclusions

The primary inclusion criterion was randomization to either GUÍA or SOC arm; participants assigned to the lead-in feedback phase were excluded (N=38). Among the randomized subjects, the primary parent/legal guardian must have completed the baseline survey (to capture baseline characteristics and potential confounders), attended the post-test GC visit, and completed the ROR1 survey to be eligible. Additionally, participants who completed the ROR1 survey more than four weeks after receipt of their genetic test results were ineligible (N=1). Applying this exclusion and inclusion criteria yielded the ROR1 analytic sample of 551 participants.

A ROR2 analytic sample was subsequently constructed. The eligibility criteria included participation in the ROR1 analytic sample and completion of the ROR2 survey. Moreover, if an amended genetic report was returned between ROR1 and ROR2 that changed the clinical interpretation of the results and required additional counseling, they were excluded from the ROR2 analytic sample (N=1). Of the 551 participants in the ROR1 analytic sample, 487 are included in the ROR2 analytic sample due to participants lost to follow-up (N=52), changes in ROR1 and ROR2 (N=1).

Primary Outcome Measures

Perceived understanding and confidence

Participants' perceived understanding of their child's genomic test results was assessed using a survey measure designed for this study that was included on the ROR1 and ROR2 surveys (see **Table S1** for question and response options). Response options were transformed for regression analysis of ROR1 data, the lowest response level (1) was omitted due to power (N=9) and levels 2 and 3 were combined, producing three analytic levels (2/3, 4, and 5). For consistency, ROR2 response options were similarly transformed resulting in individuals responding with level 1 being omitted (N=21). Missing data was not included in the final analysis (N=1 at ROR2).

Participants' confidence in explaining genetic test results was measured using a survey measure designed for this study that was asked on the ROR1 and ROR2 surveys (see **Table S1** for question and response options). There were no analytic transformations of the data, and missing data (N=3 at ROR2) was excluded from the final analysis. Partial proportional regression models were run controlling for genetic counselor, parental education, parental age, language at ROR1/ROR2 survey, child's insurance, and case-level clinical interpretation of the results.

Ordinal logistic regression was used to evaluate between arm differences in perceived understanding and confidence. When a covariate in the regression model did not satisfy the proportional odds assumption, the partial proportional odds model was used adjusting for GC, parent age (continuous), education level, language of ROR1/ROR2 surveys, child's insurance type, and case-level clinical interpretation of the results.

Secondary Outcomes Measures

Objective Understanding

The secondary outcome of participants' objective understanding of their child's GT results was measured using four questions designed for this study and asked of participants in the ROR1 and ROR2 surveys. The genetic counselor (GC) who conducted the result disclosure visit answered the same questions upon completion of the visit (see **Table S1** for questions and response options). GCs' responses were reviewed and mapped to clinical interpretation, creating a standardized "correct response" set for the individual questions. For example, the correct responses for a positive or likely positive result were "yes" to questions 1 and 2, and "no" to questions 3 and 4. GC responses were compared to this mapping, and discrepancies were queried. A GC reviewed the queries against the results, post-test note, and GUÍA (if applicable) to confirm findings and how the results were communicated to the family. The clinical team reviewed the results of the queries, and the following decisions were made: 1) change original response to standardized response (N=58 cases), or 2) leave original response and note edge case due to case complexity/nuance (N=4). A new variable was created to capture the responses that were changed. For analysis, the participant's response was transformed by comparing their

response to the GC's for each question; if the responses were the same, it was coded as 'match; if it differed, it was coded as 'no match,' producing a binary variable.

Logistic regression was conducted to analyze the four binary objective understanding variables, controlling for clinical interpretation, health system, parent's age, education, and insurance status. We did not control for genetic counselors since they were not associated with objective understanding. A summary objective understanding score was calculated as the sum of the number of matches from the four binary objective understanding questions (ranging from 0 to 4) with a higher number indicating better objective understanding. The summary score was analyzed using Poisson regression to account for the count nature of the variable. Missing data was not included in the final analysis (N=1 at ROR2).

Actionability of results and adherence to medical follow-up recommendations

Understanding of the actionability of genomic results was collected after results disclosure (ROR1) using an adapted CSER measure ("Recommended Medical Actions and Follow Through on Recommendations Attributable to Genomic Testing (MRA)"), and adherence to medical follow-up recommendations was asked at ROR2 using the CSER MRA measure (see **Table S1** for questions and response options). Due to inconsistencies in how the questions were asked between the ROR1 and ROR2 surveys, and differences in how participants interpreted and answered the survey questions, we were unable to evaluate and report on these outcomes.

Transformation of Covariates and Select Characteristics of Interest

Parent Age: Age of the parent/legal guardian was collected at baseline (**Table S2**) and calculated by subtracting the participant's self-reported date of birth from the date of baseline survey administration and was used as a continuous variable for analysis. Missing values were not included in the final analysis (N=2).

Education Level: Education level of the parent was collected at baseline (**Table S2**). For analysis, education level was collapsed into four categories: less than high school graduate (response options 1-5), high school graduate/GED, technical school, or associate degree (response options 6-9), college graduate (option 10), and college graduate plus (options 11-13). Those that selected don't know or prefer not to answer were not included in the final analysis (N=2).

Child's Insurance: Insurance status of the child was collected on the baseline survey (**Table S2**). For analysis, insurance status was binarized into "Public", which included any Government plan, and "Private", which included the private health insurance and other plans, if not government.

Population Groups: Race and ethnicity was collected at baseline (**Table S2**). For population characteristics and analysis, Hispanic/Latino(a) ethnicity was prioritized; participants who selected Hispanic/Latino(a) were re-categorized as Hispanic/Latino(a) regardless of any other race designation made. Participants that selected more than one race were re-categorized into "More than one race". All other race and ethnicity categories remained if they were the only selection made by the participant. Due to power, the three largest race and ethnicity groups were

accessed in stratification analyses: Hispanic/Latino(a) (H/L), White or European American (EA), and Black or African American (AA). Race and ethnicity was not collected for legal guardians (N=20) and were therefore excluded from the analysis.

Household Income and Number of People Supported: Poverty index was calculated using participant reported family income and number of people supported collected at baseline (**Table S2**)**.** If mean household income, accounting for the number of people supported, was at or below 200% of the 2022-2023 New York City Federal Income Guidelines, participants were categorized as living in poverty. Missing responses, 'don't know' or 'prefer not to answer' were excluded (N=63). Additionally, one participant with a reported income of '$140,000 or more' that supported 15 people was excluded as we were unable to determine whether they fell below or above the poverty level. However, 91 participants reported >$140,000 with a household size that ranged from 2 to 8 and were classified as above poverty level.

Health Literacy Level: Health literacy was captured at baseline using four survey items (**Table S2**)**.** Response options were summed to create the health literacy summary score and categorized by range of score: inadequate (4-12), marginal (13-16), and adequate (17-20). Per CSER analysis guidelines, mean imputations were calculated for those that did not provide responses for no more than half of the health literacy items (N=1).

Health System: Study research coordinators indicated the health system (Mount Sinai (MS) or Albert Einstein College of Medicine/Montefiore Medical Center (EM)) from which participants were recruited at referral (**Table S2**)**.**

Genetic Counselors: Eight study genetic counselors from MS and EM were assigned to one of the study arms; those involved in the development of GUÍA were assigned to the GUÍA arm (N=4) and the remaining assigned to the SOC arm (N=4) (**Table S2**)**.**

Survey Language Administration: The language (Spanish or English) in which surveys were administered was documented in the study database by study staff (**Table S2**). Survey language at ROR1 was used in all analyses using ROR1 data and survey language at ROR2 was used in all analyses using ROR2 data. For repeated measures analyses, language at each appropriate time point was used in the model.

Medical Interpreter Use: Genetic counselors documented whether a Spanish-speaking medical interpreter was used during the result disclosure visit (**Table S2**)**.**

Case-Level Interpretation of Genetic Test Results: The clinical interpretation of GT results was categorized by the genetic counselors as positive, likely positive, uncertain or negative based on criteria previously described in Abul-Husn et al. 2023.[1] An interpretation committee made up of study physicians with a background in medical genetics was created if discrepancies were found between the genome sequencing and targeted gene panel results. For stratified analyses, positive and likely positive were collapsed due to power, while negative and uncertain remained as they were.

Primary Indication for Testing and Neurologic Phenotype Category: The primary indication for testing was collected at recruitment from the referring provider using a study specific phenotype checklist. Providers indicated whether the primary medical concern was neurologic, cardiac, or immunologic. Participants were randomized on this selection. Participants could have more than one indication for testing. Each indication category had additional phenotype terms for further characterization. Neurologic: if epilepsy and/or intellectual developmental disability/global developmental delay; Cardiac: if congenital heart disease, cardiomyopathy, and/or cardiac arrhythmia; and Immunologic: features of immunodeficiency. Primary indication for testing data did not undergo any transformations and no data was missing. Neurologic phenotype category was analyzed for those with a primary indication of Neurologic. Due to power, further phenotype characterization of those with a primary cardiac and immunologic indication were not analyzed.

References

1. Abul-Husn, N.S., Marathe, P.N., Kelly, N.R., Bonini, K.E., Sebastin, M., Odgis, J.A., Abhyankar, A., Brown, K., Di Biase, M., Gallagher, K.M., et al. (2023). Molecular diagnostic yield of genome sequencing versus targeted gene panel testing in racially and ethnically diverse pediatric patients. Genet. Med. *25*, 100880