ARTICLE OPEN

Check for updates

Treating autism with Bumetanide: Identification of responders using Q-Finder machine learning algorithm

Hamed Rabiei[1,4], Marilyn Begnis[1,4], Eric Lemonnier[2] and Yehezkel Ben-Ari[1,3]✉

Bumetanide, a specific NKCC1 co-transporter inhibitor, restores deficient GABAergic inhibition implicated in various brain disorders, including Autism Spectrum Disorders (ASD). In keeping with this mechanism, nine successful phase 2 clinical trials, conducted by seven independent teams using an identical protocol, have shown significant improvements in ASD symptoms among individuals treated with Bumetanide. Despite these promising results, two large phase 3 clinical trials (over 400 children recruited in approximately 50 centers and covering age groups 2–6 and 7–17 years) failed with no significant difference between patients treated by placebo or Bumetanide. This failure may stem from the substantial heterogeneity of ASD symptom profiles across the study population, potentially diluting the overall observed treatment effect. To address this, we reanalyzed the phase 3 data using Q-Finder, a supervised machine learning algorithm, aiming to identify subgroups of patients who responded to the treatment. This analysis was based on clinical parameters collected at the baseline of trial and used the same standard endpoints and success criteria defined in the original phase 3 protocol. It enabled the identification of responder subgroups showing a statistically significant difference between placebo and Bumetanide treatment arms. We report detailed descriptions and statistical evaluations of these subgroups. The discovered responder subgroups, representing up to 40% of participants, were cross validated between the two study populations. These findings suggest that meaningful treatment responses can be uncovered within negative phase 3 trials, highlighting the limitations of a one-size-fits-all approach for heterogeneous conditions such as ASD. Machine learning appears to be a promising tool to support this precision medicine strategy.

Translational Psychiatry (2026)16:66 ; https://doi.org/10.1038/s41398-026-03848-3

INTRODUCTION

In a wide range of pathological conditions, the inhibitory transmitter GABA (gamma-aminobutyric acid, γ-aminobutyric acid) exerts excitatory actions due to enhanced activity of the NKCC1 cotransporter leading to high $(Cl^-)_i$ levels that shifts the polarity of GABA actions [1, 2]. Bumetanide, a NKCC1 specific inhibitor, restores GABAergic inhibition and attenuates many disorders in experimental and clinical conditions. Relying on these observations and following experimental proof of concept, we have conducted two phase 2 clinical trials, one monocentric and one multicentric with positive results, showing a significant effect of Bumetanide on symptoms of ASD [3, 4]. In parallel, Bumetanide was found to enhance visual interactions and reduce amygdala activation in constrained visual communications [5–7]. Seven additional phase 2 studies using the same protocol performed in many different countries validated these results with over 1030 children positively responding to Bumetanide [8, 9].

Following these results, two final phase 3 clinical trials were performed to evaluate the safety and efficacy of Bumetanide oral liquid formulation. These were international, randomized, double-blind, placebo-controlled studies in children and adolescents aged from 7 to 17 years (Study 1) and younger children aged from 2 to 6 years (Study 2) performed in about 50 centers across 13 countries. Trials were done for a 6-month period (Week 0 – Week 26) and patients were randomized in a 1:1 ratio in one of two parallel treatment arms: Bumetanide oral liquid formulation twice daily (BID) or placebo BID. For a comprehensive overview of the trial design, alongside detailed inclusion and exclusion criteria for participants, we refer to our previous work [10].

Although Bumetanide demonstrated significant and large effects in alleviating ASD core symptoms in several phase 2 clinical trials, the phase 3 trials did not achieve its endpoints [11]. One possible explanation is the considerable heterogeneity of ASD symptoms, potentially linked to diverse underlying etiological factors, which may dilute the observable effects of treatment at the population level. Nevertheless, clinical and statistical analyses suggest that a subset of patients did show a positive response to Bumetanide. This observation prompted us to investigate the clinical heterogeneity of children included in the trials to identify subgroups of responder patients who might respond positively to the treatment within the two phases 3 trials. Accordingly, the aim of this study is to identify patient profiles for which Bumetanide confers an above-average therapeutic benefit, using the same standard endpoints and success criteria established in the phase 3 trials [10, 11].

To this end, we relied on Q-Finder, a proprietary data mining algorithm developed by Quinten [12]. Q-Finder is a supervised machine learning tool specifically designed to identify clinically

[1]B&A Biomedical, Marseille, France. [2]Autism Expert Center and Autism Resource Center of Limousin, University Hospital Center, Limoges, France. [3]Neurochlore, Marseille, France. [4]These authors contributed equally: Hamed Rabiei, Marilyn Begnis. ✉email: ben-ari@neurochlore.fr

Received: 26 June 2025 Revised: 12 December 2025 Accepted: 20 January 2026
Published online: 03 February 2026

meaningful patient subgroups that respond differentially to treatment. It systematically explores combinations of baseline characteristics to generate interpretable rules associated with treatment response. Q-Finder has demonstrated its utility across a range of therapeutic areas for uncovering clinically meaningful patient subgroups. For instance, it was employed in the International Diabetes Management Practice Study (IDMPS) to identify subgroups of patients with type 2 diabetes showing improved glycemic control and reduced risk of hypoglycemia [12]. In another application, Q-Finder was used to detect predictive profiles associated with diabetic ketoacidosis (DKA) in type 1 diabetes, revealing both known and novel risk combinations [13]. Furthermore, in a comparative analysis of SGLT2 versus DPP-4 inhibitors, Q-Finder identified patient profiles linked to better renal function preservation with SGLT2 inhibitors [14]. These successful use-cases illustrate the potential of Q-Finder for supporting data-driven precision medicine by stratifying patient populations based on clinically relevant treatment responses or risks.

MATERIALS AND METHODS

Study populations

This analysis included participants from two Phase 3 clinical trials: Study 1, comprising subjects aged 7 to 17 years, and Study 2, comprising subjects aged 2 to 6 years. Only individuals who were randomized and received either Bumetanide or placebo were considered for inclusion. Subjects were excluded if they met any of the following criteria:

- Discontinuation from the study due to adverse events, non-medical withdrawal, or protocol violations.
- Absence of Childhood Autism Rating Scale, Second Edition (CARS2) scores at either baseline (screening day before the double-blind period, Week 0 or W0) or Week 26 (the end of the double-blind treatment period, W26).
- CARS2 baseline scores recorded after the date of randomization.

The study populations were characterized by using standard descriptive statistics to facilitate a comprehensive understanding and comparison across groups. Key variables included demographic information and baseline autism severity scores. Analyses were conducted separately for each study population to identify potential responder subgroups.

Outcome variables

The efficacy of Bumetanide compared to placebo was assessed using two outcome measures as in the original phase 3 trial: the change in total raw score of CARS2, designated as the primary endpoint, and the change in total raw score of the Social Responsiveness Scale, Second Edition (SRS-2), designated as the secondary endpoint. Both outcomes were evaluated from baseline (W0) to W26, marking the end of the double-blind treatment period. In both scales, higher scores reflect greater severity of ASD symptoms; therefore, a significant reduction from baseline is interpreted as an improvement in symptomatology (see Supplementary Figure 1).

Explanatory variables

At baseline, a wide range of variables were recorded for each participant in the Phase 3 trials. These included demographic characteristics (age, sex, country), metabolic indicators (weight, height, body mass index), laboratory parameters (biochemical, haematological, and urinary analyses), clinical examination data (vital signs), concomitant therapies and medications, medical history, and reported paediatric adverse events.

The severity of ASD symptoms was assessed using CARS2 and SRS-2. Core symptoms were evaluated according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria, and overall disorder severity was measured using the Clinical Global Impression (CGI) scale. In addition, adaptive functioning was assessed with the Vineland Adaptive Behaviour Scales, Second Edition (Vineland-II, or VABS), and cognitive abilities were evaluated through intelligence quotient (IQ) testing.

For the purposes of the present analysis, these variables were used to define patient subgroups. To enhance interpretability and focus on clinically meaningful insights, only variables with an established or potential association with ASD were retained. Specifically, this included all item-level scores from CARS2 and SRS-2, CGI ratings, and all items from VABS and DSM-5 assessments.

Feature engineering

The number of explanatory variables directly influences the number of subgroups to be analysed, and the volume of statistical tests performed, thereby increasing the risk of false positive findings. To mitigate this risk and enhance statistical power, dimensionality reduction was applied using the following methods:

CARS2 item grouping. Thirteen out of fifteen CARS2 items were aggregated into 4 composite variables based on the diagnostic domains defined in international classifications such as DSM-5 and ICD-11. This expert-guided grouping is as follows:

- Social interaction and communication disorders: average of items 1, 2, 3, 11, and 12
- Stereotyped behaviours: average of items 4 and 5
- Neurosensory disorders: average of items 7, 8, and 9
- Difficulties adapting to change: average of items 6, 10, and 13

Items 14 (cognitive level and homogeneity of intellectual functioning) and 15 (clinician's general impression of ASD severity) were excluded from this grouping due to their broader evaluative nature. This classification was developed through consensus among three clinical experts in ASD.

SRS-2 subscale aggregation and transformation. The 65 items of the SRS-2 were grouped into 5 subscales and one DSM-5-compatible domain, as defined in the standard SRS-2 Profile Sheet: Subscale 1 (social awareness), Subscale 2 (social cognition), Subscale 3 (social communication), Subscale 4 (social motivation), Subscale 5 (restrictive interest and repetitive behaviour) and Domain A (Social interaction and communication). In addition, total score, subscale scores, and DSM-5-compatible domain score were converted into T-scores to allow classification into established severity categories, while controlling for potential confounding effects of age and sex.

Confounding variables

Confounding variables are factors associated with both the explanatory and outcome variables, causing a spurious association and leading to erroneous conclusions. In the context of statistical analysis, the presence of confounding variables can bias the estimates, making it essential to account for them in the statistical models. Proper handling of these variables is pivotal to ensure the robustness and validity of the study results.

An essential consideration is the risk of overfitting due to the inclusion of an excessive number of confounding variables in the model. An overfitted model may perform well on the training data but fails to generalize effectively to new or unseen data. Given these considerations and following the analysis design of the Phase 3 trials, a balanced approach is adopted in the selection of confounders. Accordingly, several potential confounding variables were considered:

- Baseline CARS2 (or SRS-2) value: There is an anticipation that the efficacy of the treatment may vary based on the initial severity level of the disorders. By accounting for this variation, we aim to provide a more nuanced understanding of the treatment effects.
- Country of origin: It is recognized that medical practices and protocols can differ across regions. These differences may introduce variability in treatment outcomes, and therefore it was crucial to adjust for this factor in our analysis.
- Gender: Previous studies and empirical evidence suggest that gender can influence disease progression, treatment response, and other clinical outcomes. As such, it was deemed essential to account for gender-based variations when assessing the effects of the treatment.

The inclusion of these confounding variables in our statistical models ensures a more comprehensive and robust analysis, enhancing the interpretability and generalizability of our results.

Q-Finder's pipeline

Q-Finder, a proprietary subgroup discovery algorithm, is a supervised learning tool specifically designed to operate without making assumptions

about the distribution of outcome or explanatory variables. The outcomes can range from binary (e.g., 0/1), to continuous (e.g. changes in CARS2) or time-to-event outcomes. Through rigorous explorative methods, the algorithm explores the space of explanatory variables to detect regions where the outcome of interest exhibits an improved treatment effect. Thus, it can be used to identify statistically credible factors in clinical research.

The Q-Finder approach systematically delves into all variable combinations, which include both categorical and continuous types. Due to computational limitations, the algorithm has a built-in limitation of accommodating a maximum of two combined variables.

One of the primary deliverables of the Q-Finder algorithm is what we refer to as "profiles". A profile is essentially a combination of specific criteria. These criteria are paired variables-thresholds. These profiles characterize subgroups of patients who benefit from an improved treatment effect. To put it more simply, the variables constituting a profile can be continuous (defined by a range between a lower and an upper limit) or can be categorical when referring to qualitative variables.

Q-Finder's pipeline consists in 5 main steps: variable pre-processing, exhaustive subgroup generation, subgroup statistical assessment, diverse subgroup selection, and subgroup validation. The advantage of Q-Finder is that it runs an exhaustive search of subgroups to find the most optimal ones while the subgroup relevance can be validated by conventional measures of credibility (effect size, treatment effect, *p*-value etc.) and medical knowledge.

Variable pre-processing. In the preparation phase for data analysis, all numerical explanatory variables were categorized through the application of user guides and medical knowledge. So, the CARS2 domain variables were categorized based on their raw scores as follows: a score from 1 to 1.5 corresponds to the Normal category; scores from 1.5 to 2 are classified as Mildly abnormal; scores from 2.5 to 3 fall under Moderately abnormal; and scores from 3.5 to 4 are considered Severely abnormal.

Similarly, the SRS-2 subscale and domain variables were categorized based on T-scores as follows: scores below 60 were classified as Normal; scores from 60 up to 66 were categorized as Mild; scores from 66 up to 76 were considered Moderate; and scores of 76 or higher were classified as Severe.

The categorization of numerical variables offers two key advantages: first, it reduces the number of potential subgroups generated, thereby lowering the risk of false positive findings; second, it enhances the interpretability of the identified subgroups, making the clinical implications more accessible and meaningful.

Subgroup generation. A subgroup is defined by a rule, composed of 1 or 2 criteria, where each criterion represents a condition on a clinical variable. A criterion consists of a clinical variable (e.g. CGI), a comparison operator ("=" for categorical variables), and a threshold value (e.g. *Markedly ill*). The number of criteria within a rule determines its complexity (C). Subgroups of patients were systematically generated based on two levels of complexity to ensure comprehensive data interpretation:

- Complexity 1: Subgroups created using a single variable.
- Complexity 2: Subgroups created by combining two distinct variables.

For example, a rule of complexity 2 could be defined as:
(SRS-2 subscale 5 = Moderate) AND (CGI = Markedly ill)

Subjects who satisfy the rule (i.e. meet both criteria) are included in the corresponding subgroup (in-rule), while other subjects form the complementary subgroup (out-rule). A schematic illustration of this rule structure is provided in Supplementary Figure 2.

To generate rules, we relied on 26 medical variables that were recorded at baseline and known to be associated with ASD. These included categorized CARS2 total raw score and domain scores, SRS-2 total T-score and subscale and domain scores, CGI, Vineland-II domains, DSM-5 diagnostic criteria (A and B). Thresholds for each variable were derived from their established clinical scales, ensure that subgroup definitions remain both interpretable and clinically meaningful. A complete list of variables and corresponding thresholds is provided in Supplementary Table 1.

For each rule, participants with missing values in any variable involved in that rule were excluded from the corresponding subgroup analysis.

Subgroup statistical assessment. Following the generation of candidate subgroups, a multi-step assessment was conducted in which subgroups were progressively filtered based on a predefined set of credibility criteria, designed to ensure both statistical robustness and clinical relevance. These evaluation steps are outlined below:

1. **Coverage filter:** The proportion of patients within a subgroup, relative to the total study population, was required to fall between 5 and 95%. This threshold was chosen to exclude subgroups that are either too small to generalize or too large to be meaningfully distinct, thereby enhancing clinical applicability.
2. **Clinical benefit filter:** A minimum level of clinical improvement was required to ensure that identified subgroups demonstrated a meaningful response to bumetanide.
   - For CARS2 as endpoint: The effect size (ES), defined as the standardized mean difference in score changes between the bumetanide and placebo arms within the subgroup, had to be greater than or equal to 0.5. This value was selected as a conservative benchmark in the absence of a formally established minimal clinically important difference (MCID) for the CARS2 scale.
   - For SRS-2 as endpoint: The treatment effect (TE), defined as the covariate-adjusted mean difference in SRS-2 score changes between the bumetanide and placebo arms within the subgroup, was required to be greater than or equal to 10 points, consistent with the recognized MCID for the SRS-2 total score.
3. **Effect significance filter:** To ensure the statistical robustness of the findings, the *p*-value associated with the covariate-adjusted treatment effect within each identified subgroup was required to be less than 0.05 (see statistical model below).
4. **Mirror effect filter:** To confirm the specificity of the treatment effect within the identified subgroup, the complementary subgroup must not meet the previously defined clinical benefit and effect significance criteria (filters 2 and 3). This ensures that the observed benefit is distinct to the subgroup of interest and not reflective of a broader, non-specific effect across the entire population.
5. **Synergy filter**: This filter applies specifically to subgroups defined by a combination of two criteria (i.e., rules of complexity 2). To ensure that the interaction between the two variables provides added discriminatory value, the combined subgroup must demonstrate a higher effect size (for CARS2) or treatment effect (for SRS-2) than either of the individual criteria alone. This confirms that the joint presence of both conditions yields a synergistic effect that enhances the subgroup's responsiveness to bumetanide.

The treatment effect was estimated using a general linear model including the fixed categorical effects of treatment, country and gender, as well as the continuous fixed covariate of baseline score, as detailed in Section 2.5. To assess differential treatment effects across subgroups, a treatment-by-rule interaction term was incorporated into the model. This interaction term allows for the evaluation of whether the treatment effect significantly differs between the identified subgroup (in-rule) and its complement (out-rule):

$$Y = \beta_0 + \beta_1 \times \text{rule} + \beta_2 \times \text{treatment} + \beta_3 \times (\text{rule} \times \text{treatment}) + \beta_4 \times \text{Baseline} + \beta_5 \times \text{Country} + \beta_6 \times \text{Gender},$$

where

- rule = 0 for patients within the subgroup (in-rule) and 1 otherwise (out-rule).
- treatment = 0 for patients in Bumetanide arm and 1 for patients in the placebo arm.
- Y is the outcome variable (see Section 2.2).

The treatment effect is then obtained through the following coefficient once the model is fit:

- β_0: Y value for patients under treatment and in-rule
- $\beta_0 + \beta_2$: Y value for patients under placebo and in-rule

- $\beta_0 + \beta_1$: Y for patients under treatment and out-rule
- $\beta_0 + \beta_1 + \beta_2 + \beta_3$: Y for patients under placebo and out-rule
- $\beta_0 - (\beta_0 - \beta_2) = -\beta_2$: Difference in Y value between patients in-rule under treatment and placebo i.e. *treatment effect*
- $\beta_0 + \beta_1 - (\beta_0 + \beta_1 + \beta_2 + \beta_3) = -\beta_2 - \beta_3$: Difference in Y between patients out-rule under treatment and placebo.

It is important to note that during the discovery phase, all eligible subgroups were evaluated using the above general linear model. In this step, *p*-values were not corrected for multiple comparisons in order to minimize the risk of false negatives and avoid prematurely excluding potentially meaningful subgroups. Correction for multiple testing was deferred to the validation phase, where statistical significance was rigorously reassessed to control the overall false discovery rate.

Subgroup selection. A final selection was performed to obtain the most promising subgroups. Among several subgroups that share the same variables and comparison operator but different threshold values (redundant subgroups), the best subgroup in terms of ES in case of CARS2 (TE in case of SRS-2) is maintained. Some subgroups, generated by different criteria, consist of the same subjects. To ensure the diversity of subgroups, these overlapping subgroups are detected and pruned by 2 experts in ASD. Finally, the most clinically relevant subgroups are selected by the same experts.

Subgroup validation. To mitigate the risk of overfitting, subgroup validation was conducted using an independent population not involved in the initial discovery phase. Ideally, discovery and validation cohorts should exhibit comparable characteristics; however, due to constraints in recruiting a new sample, we adopted a cross-validation approach using the existing trial data. Specifically, subgroups identified in Study 1 were tested in Study 2, and vice versa, to evaluate whether they met predefined criteria for treatment response.

To limit the number of statistical tests and reduce the likelihood of false-positive findings, a two-step filtering process was implemented prior to validation. First, only subgroups covering at least 15% of the discovery population were selected. This coverage threshold was selected based on the observed distribution of subgroup sizes and practical considerations related to the feasibility of conducting future validation studies. Second, the remaining subgroups were ranked by effect size (or by treatment effect in the case of SRS-2 as the endpoint) and only the top five subgroups were retained for validation. To account for multiple comparisons during the validation step, *p*-values were adjusted using the Benjamini-Hochberg procedure, controlling the false discovery rate.

Implementation

Data extraction and Q-Finder analysis were implemented by in-house developed Python 3 scripts. We used Pandas (v.1.4.4) and Numpy (v1.23.3) libraries for data handling and visualisation. Moreover, statistical analyses, including the general linear models used for subgroup evaluation, were implemented using the StatsModels library (v.0.13.2).

RESULTS

The Q-Finder algorithm was applied separately to both study populations, Study 1 (subjects aged 7 to 17 years) and Study 2 (subjects aged 2 to 6 years), and to both the primary and secondary endpoints of the Phase 3 clinical trial: the CARS2 and SRS-2 total raw scores. This yielded a total of four independent analyses.

Population description – Study 1

Upon conducting an extensive review of all potential participants, a strict application of the study's inclusion and exclusion criteria was implemented, as delineated in Supplementary Figure 3. Out of the 261 patients initially screened for possible participation in the RCT of Study 1, a subset of 211 patients met all eligibility requirements. Of this eligible cohort, 178 patients were successfully randomized, received the designated treatment, completed the entirety of the study, and had both initial and concluding CARS2 values recorded ($N = 86$ in Bumetanide arm and $N = 92$ for placebo arm).

As illustrated in Supplementary Table 2, the cohort of 178 patients who are included in our analysis predominantly consists of younger patients with a mean age of 10.4 years (SD = 3.0) and 10.3 years (2.7) in Bumetanide and placebo arms respectively. Both arms had higher prevalence of male (80.2 and 82.6%) and Spanish patients (20 and 22%). The baseline average CARS2 total raw scores were 40.2 (5.0) and 39.8 (4.8) points in Bumetanide and placebo arms, respectively. They decreased to 36.8 (6.0) and 36.8 (5.5) points in Week 26. Consequently, in case of CARS2 as endpoint, the difference between Bumetanide and placebo arms is not significant ($p = 0.452$, effect size = 0.09). The baseline average SRS-2 scores were 116.6 (23.6) and 114.7 (23.0) points in Bumetanide and placebo arms, respectively. They decreased to 101.2 (26.4) and 102.0 (25.5) points in Week 26. Consequently, in case of SRS-2 as endpoint, the difference between Bumetanide and placebo arms is not significant ($p = 0.617$, treatment effect =1.58).

Population description – Study 2

Out of the 274 patients initially screened for possible participation in the RCT, a subset of 211 patients met all eligibility requirements. Of this eligible cohort, 181 patients were successfully randomized, received the designated treatment, completed the entirety of the study, and had both initial and concluding CARS2 values recorded ($N = 89$ in Bumetanide arm and $N = 92$ for placebo arm; see Supplementary Figure 3).

As illustrated in Supplementary Table 2, patient age was uniformly distributed, yielding an average age of 4.3 (1.2) and 4.6 (1.2) in Bumetanide and placebo arms respectively. Both arms had higher prevalence of male (84.3 and 82.6%). Polish (16%) and Spanish (16%) patients had higher prevalence in Bumetanide and placebo arms respectively. The baseline average CARS2 total raw scores were 41.3 (4.9) and 40.9 (5.6) points in Bumetanide and placebo arms, respectively. They decreased to 37.7 (6.5) and 37.0 (6.4) points in Week 26. Consequently, in case of CARS2 as endpoint, the difference between Bumetanide and placebo arms is not significant ($p = 0.723$, effect size $= -0.06$). The baseline average SRS-2 scores were 113.5 (26.1) and 115.9 (21.3) points in Bumetanide and placebo arms, respectively. They decreased to 97.4 (28.1) and 97.2 (27.8) points in Week 26. Consequently, in case of SRS-2 as endpoint, the difference between Bumetanide and placebo arms is not significant ($p = 0.907$, effect size $= -0.38$).

Identification of responders – Study 1

The Q-Finder algorithm generated 2415 candidate subgroups in the Study 1 dataset using combinations of up to two criteria (see Fig. 1). Among these, 1034 subgroups met the predefined coverage criterion, with representation between 5 and 95% of the total population. By fitting a generalized linear model with CARS2 total raw score as endpoint, 9 subgroups demonstrated a statistically significant difference between Bumetanide and placebo arms (ES ≥ 0.5, and *p*-value < 0.05). These subgroups were further retained based on synergy and redundancy filters while accounting for diversity. The final set of 9 discovered subgroups, ranked by coverage, is presented in Table 1.

The largest discovered subgroup achieved a maximum coverage of 40% ($N = 71$). This subgroup was characterized by patients whose CARS2 neurosensory domain was rated as *moderately abnormal* and whose SRS-2 subscale 3 (social communication) was rated as *severe*. These patients demonstrated a potential response to bumetanide, with an effect size of 0.53 and a corresponding *p*-value of 0.039, indicating a statistically significant and clinically relevant treatment effect.

The subgroup with the highest observed effect size consisted of patients whose CARS2 neurosensory domain was rated as *mildly abnormal* and who responded *"No"* to item B3 of the DSM-5 criteria (i.e. highly restricted, fixated interests that are abnormal in intensity or focus). This subgroup demonstrated a large clinical

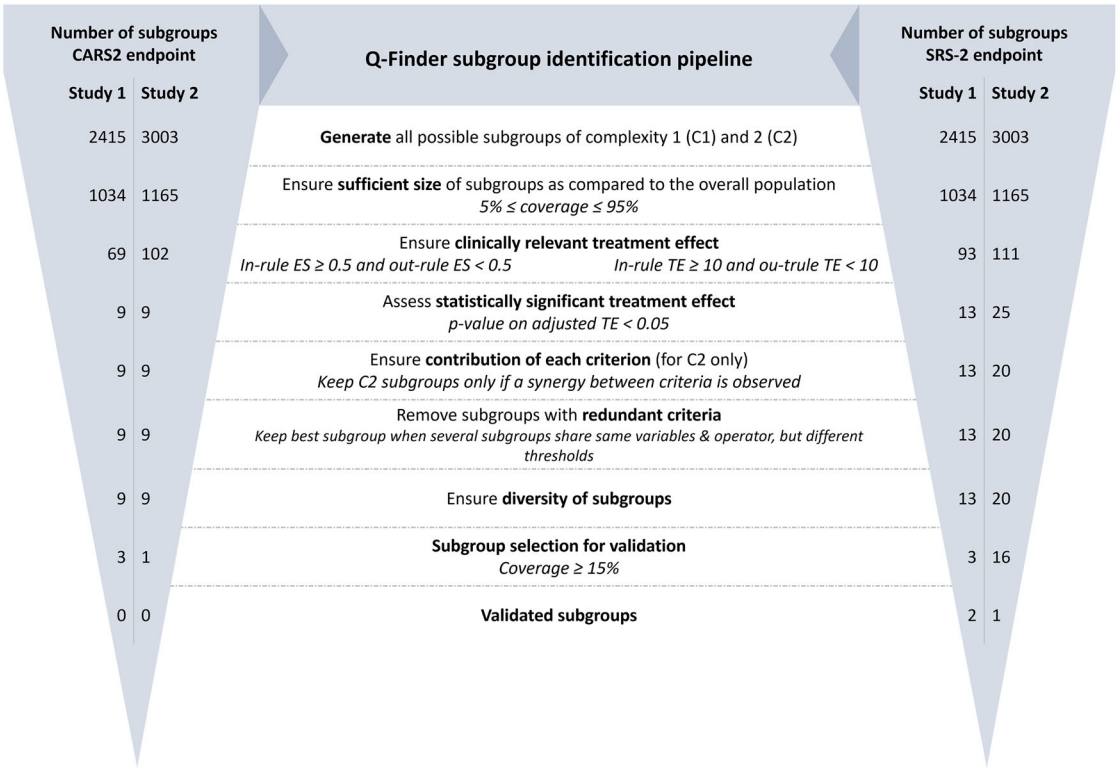

Fig. 1 Overview of the Q-Finder algorithm pipeline used to identify responder subgroups to bumetanide treatment in Study 1 and Study 2 populations using the CARS2 total raw score and SRS-2 total raw score as the primary and secondary endpoints. ES effect size, TE treatment effect.

Table 1. Responder subgroups identified in the Study 1 population using CARS2 total raw score as the primary efficacy endpoint.

Subgroup description	Coverage N (%)	*p*-value	ES
(CARS2 neurosensory domain = Moderately abnormal) AND (SRS-2 subscale 3 = Severe)	71 (40)	0.039	0.53
(SRS-2 subscale 4 = Severe) AND (DSM5-B severity = 2)	32 (18)	0.029	0.58
(SRS-2 subscale 4 = Severe) AND (DSM5-A severity = 2)	27 (15)	0.04	0.51
(CARS2 social interaction and communication domain = Severely abnormal) AND (CARS2 stereotyped behavior domain = Moderately Abnormal)	21 (12)	0.015	0.93
(CARS2 social interaction and communication domain = Severely Abnormal) AND (SRS-2 subscale 4 = Moderate)	20 (11)	0.017	1.29
(CARS2 social interaction and communication domain = Severely Abnormal) AND (DSM5-A severity = 2)	14 (8)	0.006	0.85
(CARS2 stereotyped behavior domain = Severely Abnormal) AND (SRS-2 subscale 1 = Moderate)	10 (6)	0.03	0.97
(SRS-2 subscale 3 = Severe) AND (DSM5-B1 = No)	10 (6)	0.036	1.24
(CARS2 neurosensory domain = Mildly abnormal) AND (DSM5-B3 = No)	9 (5)	0.026	1.46

Each entry includes the defining criteria, coverage within the population, effect size (ES), and associated p-value.

benefit, with an effect size of 1.46 and a *p*-value of 0.026, indicating statistical significance. Although this subgroup represented a relatively small portion of the study population (5%, $N = 9$), the magnitude of the observed effect suggests a potentially meaningful therapeutic response to bumetanide.

Considering SRS-2 as endpoint (see Fig. 1), the generalized linear model identified 13 subgroups in which the difference between the Bumetanide and placebo arms was statistically significant, meeting the criteria of an adjusted treatment effect (TE) ≥ 10 and *p*-value < 0.05. These subgroups were subsequently retained after applying synergy, redundancy, and diversity filters to ensure meaningful, non-overlapping, and clinically interpretable findings. The final set of 13 subgroups is presented in Table 2.

The subgroup with the highest coverage, encompassing 30% of the Study 1 population ($N = 54$), was characterized by patients whose CARS2 domain for adaptation to environmental changes was rated as *mildly abnormal* and who responded *"Yes"* to item B1 of the DSM-5 (i.e. stereotyped or repetitive motor movements, use of objects, or speech). This subgroup demonstrated a statistically significant adjusted treatment effect of 12.71, with a corresponding *p*-value of 0.031, indicating a meaningful differential response to bumetanide.

In terms of treatment effect magnitude, the most prominent subgroup comprised patients whose CARS2 domain for social interaction and communication was rated as *severely abnormal*, and whose adaptation to environmental changes was rated as *mildly abnormal*. This subgroup exhibited a substantial adjusted

Table 2. Responder subgroups identified in the Study 1 population using SRS-2 total raw score as the secondary efficacy endpoint.

Subgroup description	Coverage N (%)	*p*-value	TE
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5-B1 = Yes)	54 (30)	0.031	12.71
(CARS2 Adaptation to environmental changes domain = Mildly abnormal) AND (Baseline SRS-2 total T-score = Severe)	45 (25)	0.044	13.04
(Baseline CARS2 total raw score = Severely abnormal) AND (CARS2 Adaptation to environmental changes domain = Mildly abnormal)	38 (21)	0.011	17.81
(CARS2 Adaptation to environmental changes domain = Mildly abnormal) AND (CGI = Markedly ill)	25 (14)	0.04	18.72
(SRS-2 subscale 2 = Moderate) AND (Baseline SRS-2 total T-score = Severe)	24 (13)	0.036	17.84
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (SRS-2 subscale 1 = Severe)	23 (13)	0.021	20.71
(Baseline CARS2 total raw score = Severely abnormal) AND (DSM5-B3 = No)	22 (12)	0.017	23.00
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5-B severity = 3)	17 (10)	0.038	21.14
(CARS2 stereotyped behavior domain = Severely abnormal) AND (CARS2 adaptation to environmental changes domain = Mildly Abnormal)	14 (8)	0.041	24.51
(CARS2 social interaction and communication domain = Severely abnormal) AND (CARS2 adaptation to environmental changes domain = Mildly abnormal)	13 (7)	0.006	33.90
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5-B3 = No)	13 (7)	0.014	29.40
(SRS-2 subscale 4 == Mild) AND (DSM5-A severity = 3)	13 (7)	0.043	24.44
(SRS-2 subscale 1 = Mild) AND (SRS-2 subscale 3 = Moderate)	12 (7)	0.026	27.33

Each entry includes the defining criteria, coverage within the population, treatment effect (TE), and associated p-value.

Table 3. Responder subgroups identified in the Study 2 population using CARS2 total raw score as the primary efficacy endpoint.

Subgroup description	Coverage N (%)	*p*-value	ES
(Baseline SRS-2 total T-score = Moderate) AND (DSM5-B2 = Yes)	33 (18)	0.041	0.72
(SRS-2 social interaction and communication domain = Moderate) AND (DSM5-B2 = Yes)	26 (14)	0.016	1.02
(SRS-2 subscale 5 = Moderate) AND (CGI = Markedly ill)	20 (11)	0.029	1.10
(SRS-2 subscale 4 = Moderate) AND (Baseline SRS-2 total T-score = Moderate)	18 (10)	0.032	0.74
(SRS-2 subscale 5 = Moderate) AND (SRS-2 social interaction and communication domain = Moderate)	17 (9)	0.014	1.24
(CARS2 stereotyped behavior domain = Severely abnormal) AND (DSM5-A severity = 2)	15 (8)	0.03	1.92
(CARS2 neurosensory domain = Moderately abnormal) AND (DSM5-B4 = No)	15 (8)	0.03	0.72
(SRS-2 subscale 4 = Moderate) AND (SRS-2 social interaction and communication domain = Moderate)	12 (7)	0.012	1.25
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (SRS-2 subscale 4 = Mild)	9 (5)	0.013	3.10

Each entry includes the defining criteria, coverage within the population, effect size (ES), and associated p-value.

treatment effect of 33.9, with a *p*-value of 0.006, indicating strong statistical significance. Despite its relatively limited coverage of 7% ($N = 13$), the observed effect suggests a highly favorable response to bumetanide in this specific clinical profile.

Identification of responders – Study 2

In Study 2, the Q-Finder algorithm generated 3003 subgroups based on combinations of up to two criteria (see Fig. 1). Among these, 1165 subgroups met the coverage threshold (5% ≤ coverage ≤ 95%) to ensure adequate population representation. Applying a generalized linear model with CARS2 total raw score as the endpoint, 9 subgroups showed a statistically significant difference between Bumetanide and placebo arms (effect size ≥ 0.5 and p-value < 0.05). After applying synergy, redundancy, and diversity filters, 9 subgroups were retained as credible responders. These final subgroups, ranked by coverage, are detailed in Table 3.

The subgroup with the highest coverage, representing 18% of the Study 2 population ($N = 33$), included patients with a baseline total SRS-2 T-score categorized as *moderate* and who responded *"Yes"* to item B2 of the DSM-5 (i.e. insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal nonverbal behavior). This subgroup demonstrated an effect size of 0.72 with a *p*-value of 0.041, suggesting a meaningful treatment response to bumetanide.

In terms of effect size magnitude, the most notable subgroup comprised patients whose CARS2 domain for adaptation to environmental changes was rated as *mildly abnormal*, and whose SRS-2 subscale 4 (social motivation) was also evaluated as *mild*. This subgroup showed a substantial effect size of 3.1 with a *p*-value of 0.013, despite representing only 5% of the population ($N = 9$), indicating a potentially high benefit for a narrowly defined patient profile.

Considering SRS-2 as endpoint (see Fig. 1), the generalized linear model identified 25 subgroups within the Study 2 population that demonstrated a statistically significant difference between Bumetanide and placebo arms, defined by an adjusted treatment effect ≥ 10 and p-value < 0.05. Following the application of synergy, redundancy, and diversity filters, 20 subgroups were retained as credible responder profiles, shown in Table 4.

The subgroup with the highest coverage, comprising 36% of the population ($N = 66$), included patients with the CARS2 adaptation to environmental changes domain rated as *mildly abnormal*. This subgroup exhibited a statistically significant adjusted treatment effect of 12.56 ($p = 0.019$).

Table 4. Responder subgroups identified in the Study 2 population using SRS-2 total raw score as the secondary efficacy endpoint.

Subgroup description	Coverage N (%)	*p*-value	TE
(CARS2 adaptation to environmental changes domain = Mildly abnormal)	66 (36)	0.019	12.56
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (Vineland socialization domain = Low)	61 (34)	0.008	14.89
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5-B1 = Yes)	61 (34)	0.007	14.96
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5-B4 = Yes)	61 (34)	0.004	16.17
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (Vineland communication domain = Low)	60 (33)	0.006	15.70
(CARS2 neurosensory domain = Moderately abnormal) AND (CARS2 adaptation to environmental changes domain = Mildly abnormal)	54 (30)	0.025	13.27
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (SRS-2 subscale 2 = Severe)	48 (27)	0.006	16.58
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (Baseline SRS-2 total T-score = Severe)	48 (27)	0.007	16.68
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (SRS-2 social interaction and communication domain = Severe)	47 (26)	0.011	15.57
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5A severity = 3)	47 (26)	0.008	17.10
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (SRS-2 subscale 1 = Severe)	43 (24)	0.031	14.35
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (SRS-2 subscale 3 = Severe)	43 (24)	0.004	18.24
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5-B severity = 3)	36 (20)	0.023	16.80
(SRS-2 subscale 3 = Severe) AND (SRS-2 subscale 4 = Moderate)	36 (20)	0.016	17.21
(CARS2 stereotyped behavior domain = Moderately abnormal) AND (SRS-2 subscale 1 = Moderate)	29 (16)	0.045	16.82
(CARS2 social interaction and communication domain = Severely abnormal) AND (CARS2 adaptation to environmental changes domain = Mildly abnormal)	28 (15)	0.027	17.87
(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (CGI = Severely ill)	26 (14)	0.017	20.48
(SRS-2 subscale 1 = Moderate) AND (DSM5-B severity = 2)	20 (11)	0.042	21.46
(CARS2 social interaction and communication domain = Severely abnormal) AND (SRS-2 subscale 1 = Moderate)	15 (8)	0.039	23.25
(CARS2 stereotyped behavior domain = Moderately abnormal) AND (SRS-2 subscale 5 = Normal)	9 (5)	0.03	38.02

Each entry includes the defining criteria, coverage within the population, treatment effect (TE), and associated p-value.

In terms of magnitude of treatment effect, the most notable subgroup consisted of patients with the CARS2 stereotyped behavior domain rated as *moderately abnormal* and SRS-2 subscale 5 (restrictive interest and repetitive behaviour) evaluated as *normal*. Despite covering only 5% of the population ($N = 9$), this subgroup demonstrated a pronounced adjusted TE of 38.02 ($p = 0.03$), indicating a strong response to bumetanide.

Validation

Using CARS2 as the primary endpoint, 3 of 9 subgroups identified in the Study 1 population and 1 of 9 subgroup identified in the Study 2 population met the minimum coverage threshold (≥15%) for validation. However, none of these subgroups showed a statistically significant difference between the bumetanide and placebo arms in the corresponding validation population ($p \geq 0.05$); see Fig. 1.

In contrast, when using SRS-2 as the secondary endpoint, 13 subgroups were initially discovered in the Study 1 population. Of these, 3 subgroups met the coverage criterion and proceeded to validation; 2 subgroups were successfully validated as responders, showing a significant treatment effect. In the Study 2 population, 20 subgroups were discovered, with 16 meeting the coverage threshold. Among the top 5 subgroups ranked by treatment effect, 1 subgroup was validated as responder in the corresponding validation population. The validated subgroups from both studies are presented in Table 5.

DISCUSSION

Large-scale clinical trials targeting children with ASD and other neurodevelopmental disorders have consistently failed to demonstrate significant treatment effects. These failures underscore the substantial heterogeneity of symptoms, suggesting that a single therapeutic approach is unlikely to be effective across the entire spectrum. Consequently, identifying subpopulations of individuals who are more likely to respond to specific treatments has become a key objective in the field. This concept, central to precision medicine, is increasingly recognized as critical for designing successful clinical trials.

The present study illustrates one such approach by applying a machine learning algorithm, Q-Finder, to reanalyze data from a phase 3 clinical trial of Bumetanide. Our analysis focused on clinical variables that are available at baseline and are directly relevant to the core and associated symptoms of autism. This targeted preselection enhances the interpretability of the identified subgroups and facilitates the design and implementation of follow-up studies. Alternative possibilities of stratification notably based on genetics have limited applications. To date, only a small subset of ASD cases can be attributed to single-gene mutations or copy-number variants. Moreover, even in genetically defined subgroups such as Fragile X syndrome (FXS), which has a clearer etiology, clinical trials have shown mixed results. For instance, early phase trials in FXS using targeted agents (e.g., mGluR5 antagonists) failed to demonstrate consistent efficacy,

Table 5. Statistical results of validated responder subgroups from both studies using SRS-2 as the efficacy endpoint.

Population	Subgroup description	Coverage %	*p*-value	TE
Study 1 / Study 2	(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (DSM5-B1 = Yes)	30 / 34	0.011	12.71 / 14.96
Study 1 / Study 2	(CARS2 adaptation to environmental changes domain = Mildly abnormal) AND (Baseline SRS-2 total T-score = Severe)	25 / 27	0.011	13.04 / 16.68
Study 2 / Study 1	(CARS2 social interaction and communication domain = Severely abnormal) AND (CARS2 adaptation to environmental changes domain = Mildly abnormal)	15 / 7	0.032	17.87 / 33.90

For each subgroup, the population coverage and treatment effect (TE) are reported for both discovery / validation phases. p-values in the validation phase have been adjusted for multiple comparisons using the Benjamini-Hochberg procedure.

and marked variability in treatment response was noted despite common genetic status [15].

More recently, neurophysiological approaches, particularly resting-state EEG, have emerged as promising tools for identifying bumetanide responders. In the *Bumetanide in Autism Medication and Biomarker (BAMBI)* study [16], focusing on the Repetitive Behavior Scale–Revised (RBS-R) as the main endpoint, bumetanide treatment was associated with increased alpha-band power and improved functional excitation–inhibition balance, changes that correlated with clinical improvements in repetitive behaviors and social responsiveness. Notably, EEG-clinical correlations were confined to subgroups showing medium to high RBS-R improvement, suggesting that these neurophysiological signatures reflect meaningful treatment responsiveness. Moreover, machine learning models trained on baseline RBS-R scores and EEG features could predict medium and high responders with 80 and 92% accuracy, respectively. These findings underscore the high potential of EEG-based biomarkers to complement clinical and data-driven methods such as Q-Finder, collectively advancing precision medicine approaches in ASD.

Through our approach, we identified a range of responder subgroups, some comprising up to 40% of the original phase 3 trial population, and others characterized by a strong treatment effect. Using the SRS-2 total raw score as endpoint, two subgroups from Study 1 and one from Study 2 were successfully cross-validated. Notably, all validated subgroups shared a common feature: a mildly abnormal rating in the CARS2 "adaptation to environmental changes" domain. This criterion, in combination with others, appears to define subpopulations responsive to Bumetanide across both trial cohorts (Table 5). From a statistical perspective, these subgroups represent credible candidates for further validation and potentially for targeted clinical trials.

The results of this study can inform the design of future clinical trials investigating bumetanide treatment in children with ASD. Specifically, once a responder subgroup has been identified and validated, its defining characteristics can be incorporated into the inclusion criteria of subsequent studies. For example, if the first validated subgroup is selected (see Table 5), the two defining rules should be applied as inclusion criteria:

- CARS2 "Adaptation to environmental changes" domain = Mildly abnormal
- DSM-5 B1 = Yes

This study has several limitations. First, to mitigate the risk of over fitting and false positive discovery, we limited the number of explanatory variables by grouping conceptually related CARS2 items into broader domains, following the diagnostic categories defined in well-known classifications such as the DSM-5 and ICD-11. Although this aggregation was guided by clinical reasoning and validated through expert consensus among three clinicians, it may not align with standard practice and could present interpretability challenges for some practitioners. Secondly, subgroup discovery inherently involves many statistical tests, which increases the risk of overfitting and false positive findings. To address this issue, we adapted several strategies, including the preselection of relevant explanatory and confounding variables, targeted feature engineering, and variable preprocessing. These steps helped reduce the dimensionality of the data and limit the number of subgroups generated and tests to be performed. Additionally, we performed cross-validation of the discovered subgroups in an independent cohort to assess the generalizability of the results. However, while the discovery and validation populations share several characteristics such as the severity of ASD symptoms, they also differ in key aspects, most notably the age range. Therefore, the findings should be interpreted with caution until further validated in fully independent cohorts with comparable demographic and clinical profiles.

In conclusion, our findings highlight the inherent limitations of applying a single therapeutic approach to a highly heterogeneous condition such as ASD. To improve the likelihood of success in large clinical trials, it is essential to define clinically meaningful subpopulations of patients who are more likely to benefit from specific treatments. While this stratified approach introduces complexity, it represents a necessary step toward personalized intervention strategies in ASD. In the absence of validated biological markers capable of identifying responder subgroups at scale, data-driven methodologies, particularly those based on machine learning, offer a promising and practical alternative by leveraging accessible clinical parameters. It should be noted that our analysis does not directly assess the effect of bumetanide on specific symptom domains, but the identification of responder subgroups characterized by core autism symptoms may help guide future investigations into the potential mechanisms of action of the treatment, as well as the specific symptom domains in which its clinical effects may be most pronounced. Machine learning represents a promising tool to uncover novel research directions and to support the development of precision medicine approaches.

DATA AVAILABILITY

Data will be made available on request.

REFERENCES

1. Ben-Ari Y. NKCC1 chloride importer antagonists attenuate many neurological and psychiatric disorders. Trends Neurosci. 2017;40:536–54.
2. Savardi A, Borgogno M, De Vivo M, Cancedda L. Pharmacological tools to target NKCC1 in brain disorders. Trends Pharmacol Sci. 2021;42:1009–34.
3. Lemonnier E, Degrez C, Phelep M, Tyzio R, Josse F, Grandgeorge M, et al. A randomised controlled trial of bumetanide in the treatment of autism in children. Transl Psychiatry. 2012;2:e202–e202.
4. Lemonnier E, Villeneuve N, Sonie S, Serret S, Rosier A, Roue M, et al. Effects of bumetanide on neurobehavioral function in children and adolescents with autism spectrum disorders. Transl Psychiatry. 2017;7:e1056–9.
5. Hadjikhani N, Zürcher NR, Rogier O, Ruest T, Hippolyte L, Ben-Ari Y, et al. Improving emotional face perception in autism with diuretic bumetanide: a proof-of-concept behavioral and functional brain imaging pilot study. Autism. 2015;19:149–57.
6. Hadjikhani N, Åsberg Johnels J, Zürcher NR, Lassalle A, Guillon Q, Hippolyte L, et al. Look me in the eyes: constraining gaze in the eye-region provokes abnormally high subcortical activation in autism. Scientific Reports. 2017;7:1–7.
7. Hadjikhani N, Åsberg Johnels J, Lassalle A, Zürcher NR, Hippolyte L, Gillberg C, et al. Bumetanide for autism: More eye contact, less amygdala activation. Sci Rep. 2018;8:1–8.
8. Wang T, Shan L, Miao C, Xu Z, Jia F Treatment effect of bumetanide in children with autism spectrum disorder: A systematic review and meta-analysis. Front Psychiatry 2021;12. https://doi.org/10.3389/FPSYT.2021.751575.
9. Xiao HL, Zhu H, Jing JQ, Jia SJ, Yu SH, Yang CJ. Can bumetanide be a miraculous medicine for autism spectrum disorder: Meta-analysis evidence from randomized controlled trials. Res Autism Spectr Disord. 2024;114:102363.
10. Crutel V, Lambert E, Penelaud PF, Albarrán Severo C, Fuentes J, Rosier A, et al. Bumetanide oral liquid formulation for the treatment of children and adolescents with autism spectrum disorder: design of two phase III studies (SIGN Trials). J Autism Dev Disord. 2021;51:2959–72.
11. Fuentes J, Parellada M, Georgoula C, Oliveira G, Marret S, Crutel V, et al. Bumetanide oral solution for the treatment of children and adolescents with autism spectrum disorder: Results from two randomized phase III studies. Autism Research. 2023;16:2021–34.
12. Esnault C, Gadonna ML, Queyrel M, Templier A, Zucker JD. Q-Finder: An algorithm for credible subgroup discovery in clinical data analysis — an application to the international diabetes management practice study. Front Artif Intell. 2020;3:559927 https://doi.org/10.3389/FRAI.2020.559927/FULL.
13. Ibald-Mulli A, Seufert J, Grimsmann JM, Laimer M, Bramlage P, Civet A, et al. Identification of predictive factors of diabetic ketoacidosis in type 1 diabetes using a subgroup discovery algorithm. Diabetes Obes Metab. 2023;25:1823–9.
14. Zhou FL, Watada H, Tajima Y, Berthelot M, Kang D, Esnault C, et al. Identification of subgroups of patients with type 2 diabetes with differences in renal function preservation, comparing patients receiving sodium-glucose co-transporter-2 inhibitors with those receiving dipeptidyl peptidase-4 inhibitors, using a supervised machine-learning algorithm (PROFILE study): A retrospective analysis of a Japanese commercial medical database. Diabetes Obes Metab. 2019;21:1925.
15. Jacquemont S, Berry-Kravis E, Hagerman R, Von Raison F, Gasparini F, Apostol G, et al. The challenges of clinical trials in fragile X syndrome. Psychopharmacology (Berl). 2014;231:1237–50.
16. Juarez-Martinez EL, Sprengers JJ, Cristian G, Oranje B, van Andel DM, Avramiea AE et al. Prediction of behavioral improvement through resting-state electroencephalography and clinical severity in a randomized controlled trial testing bumetanide in autism spectrum disorder. Biol Psychiatry Cogn Neurosci Neuroimaging 2021. https://doi.org/10.1016/J.BPSC.2021.08.009.

AUTHOR CONTRIBUTIONS

HR and MB have developed the Machine Learning program and contributed to paper writing & submission. Y B-A has supervised the project and written the paper.

COMPETING INTERESTS

The study was funded by Neurochlore and BA-biomedical: 2 startups dedicated to treat Autism (neurochlore) and use IA and Machine Learning in this aim. Y Ben-Ari is CEO and shareholder of both startups. H Rabiei and M Begnis are paid by BA-Biomedical but are not shareholders. E Lemonnier has no conflict of interest.

ETHICAL COMMITTEE AND PARENT CONSENTS

This paper is based on a re-analysis of a large multicentric phase 3 trial made by Servier (2021-2023) in 40 centers (Europe, Brazil, Australia) and all ethical issues and other requirements were met then and published (Crutel et al Journal of autism and neurodevelopmental disorders 2021 & Fuentes et al -Autism research 2023).

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-026-03848-3.

Correspondence and requests for materials should be addressed to Yehezkel Ben-Ari.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.