Heterogeneity in response to repeated intranasal oxytocin in schizophrenia and autism spectrum disorders: A meta-analysis of variance

Daniel Martins | Maria Paduraru | Yannis Paloyelis

Intranasal oxytocin (OT) has been suggested as a putative adjunctive treatment for patients with schizophrenia and autism spectrum disorders (ASD). Here, we examine available evidence from trials investigating the effects of repeated administrations of intranasal OT on the core symptoms of patients with schizophrenia and ASD, focusing on its therapeutic efficacy and heterogeneity of response (meta-ANOVA). Repeated administration of intranasal OT does not improve most of the core symptoms of schizophrenia and ASD, beyond a small tentative effect on schizophrenia general symptoms. However, we found significant moderator effects for dose in schizophrenia total psychopathology and positive symptoms, and percentage of included men and duration of treatment in schizophrenia general symptoms. We found evidence of heterogeneity (increased variance) in the response of schizophrenia negative symptoms to intranasal OT compared with placebo, suggesting that subgroups of responsive and non-responsive patients might coexist. For other core symptoms of schizophrenia, or any of the core symptom dimensions in ASD, the response to repeated treatment with intranasal OT did not show evidence of heterogeneity.

KEYWORDS
autism spectrum disorders, coefficient of variation ratio, efficacy, heterogeneity, intranasal oxytocin, schizophrenia

1 | INTRODUCTION

Over the last two decades, there has been enthusiasm for the possibility of using intranasal oxytocin (OT) as an adjunctive therapeutic in many psychiatric disorders (Quintana et al., 2021). This enthusiasm has been largely motivated by the potential OT holds to improve psychopathology in multiple domains that existing therapeutic agents cannot fully target, such as social functioning (Young, 2015).

A number of small scale randomized clinical trials (RCT) have evaluated the effects of regimens involving the repeated administration of intranasal OT as a mono- or adjunctive therapy in patients with psychiatric disorders (Cochran et al., 2013). Most of these studies have focused on patients with schizophrenia (Zheng et al., 2019) and autism spectrum disorders (ASD) (Ooi et al., 2017). So far, these RCTs have largely failed to provide convincing evidence that treatment regimens involving the repeated administration of intranasal OT improve any of the core symptoms of patients with schizophrenia (Williams & Burkner, 2017; Zheng et al., 2019) or ASD (Ooi et al., 2017; Wang et al., 2019). Some evidence contradicting all previous meta-analyses was reported in a recent meta-analysis including nine schizophrenia...
and 10 ASD RCTs published up to 2019: Peled-Avron et al. found that repeated intranasal OT improved total psychopathology in patients with schizophrenia, and repetitive behaviours and social functioning (but not global symptoms) in patients with ASD (Peled-Avron et al., 2020). However, this last meta-analysis did not control for pretreatment symptom scores, which may have led to biased estimates of treatment efficacy (i.e. differences in symptoms severity between the placebo and intranasal OT groups during the post-treatment assessment might simply reflect pre-existing differences between groups at baseline).

A further shortcoming of existing meta-analyses is that they rely on mean group level estimates of efficacy and, therefore, cannot elucidate whether a repeated intranasal OT treatment regime might have beneficial effects for a subset of responsive patients. Such heterogeneity in treatment response is often evoked to explain the lack of overall findings at group-level analyses, as the effects on a subset of responsive patients might be diluted by the lack of effects on non-responsive patients (Bradley & Woolley, 2017; Cochrane et al., 2013), but we are not aware of any study testing this hypothesis formally. The idea of substantial underlying heterogeneity in the response to repeated intranasal OT treatment in patients with schizophrenia and ASD is supported by two sources of evidence. First, both schizophrenia (Liang & Greenwood, 2015) and ASD (Mottron & Bzdok, 2020) are highly heterogeneous disorders; thus, it is plausible that intranasal OT might benefit specific subsets of patients where a deficit in OT signaling might contribute to symptoms to a greater extent. Second, a number of inter-individual factors have also been shown to moderate the response to intranasal OT, such as (epi)genetic variation (Chen et al., 2020; Feng et al., 2015) and childhood trauma (Joseph et al., 2019; van Zuiden et al., 2017), likely through effecting on the capacity of the OT receptor to respond to OT binding (Kraaijenvanger et al., 2019). Hence, to properly evaluate the therapeutic potential of a daily-repeated intranasal OT treatment regime in patients with schizophrenia and ASD, it is important to investigate whether subgroups of patients may show a differential response to treatment that is masked by mean group level analyses.

The aims of this study are twofold. First, given the inconsistencies and methodological shortcomings among recent meta-analysis, and the emergence of new RCTs that have not been previously included (Bernaerts et al., 2020), we conducted updated meta-analyses of all RCTs investigating the effects of repeated intranasal OT treatment on the core symptoms of patients with schizophrenia and ASD, using gold-standard methods. Specifically, to estimate the effects of repeated intranasal OT on symptoms we used the standardized mean change using raw score standardization (SMCR) statistic (Morris, 2000), which unlike other effect sizes such as Hedge’s g, not only contrasts treatment and control groups, but also controls for possible differences in the pretreatment values. This approach would help us to re-evaluate evidence regarding the efficacy of intranasal OT in these two populations of patients using state-of-the-art methodology that protects researchers and clinicians from over-optimistic and perhaps misguided conclusions. Second, we investigated whether heterogeneity (i.e. subsets of responsive and non-responsive) might exist in the response to repeated intranasal OT treatment by conducting meta-analyses of variance. Recent meta-analyses using this approach investigated indirectly the presence of individual responses in RCTs (Mizuno et al., 2020; Radua et al., 2020; Senior et al., 2016; Winkelbeiner et al., 2019). These methods are based on the quantification of variance in the treatment group compared with the placebo group and have already been applied to address many other important questions related to heterogeneity in psychiatry (Brugger et al., 2020; Mizuno et al., 2020; Pillinger et al., 2019; Radua et al., 2020; Winkelbeiner et al., 2019). A key prediction of this variance model is that, in the presence of treatment-by-individual (or groups of individuals) interactions, variance in the response to treatment should be higher in the intranasal OT than in the placebo group. In this case, group-level estimates are unlikely to be good approximations of the estimates for the individual, calling for the need of a personalized or stratified medicine approach.

2 | METHODS

2.1 | Literature search

We followed the PRISMA guidelines (Moher et al., 2010) (flowcharts in Figures S1 and S2) during our systematic review process. We conducted searches in the PubMed, Web of Science and Cochrane Central Register of Controlled Trials from inception until 18th June 2020. To search for intranasal OT schizophrenia RCTs, we used the following search query: ‘(Oxytocin OR Pitocin) AND (Schizo* OR Dementia praecox OR Psychosis OR Psychotic)’. To search for intranasal OT ASD RCTs, we used the following search query: ‘(Oxytocin OR Pitocin) AND (Autism OR ASD OR Autist* OR Asperger OR Pervasive developmental disorder)’. Our search queries were adapted in accordance with the specification of each database. Our search strategy was tailored to capture all OT studies in schizophrenia and ASD and not only RCTs. We used a broad search query to minimize search bias.

2.2 | Study selection

We performed study selection with the help of the reference manager software Rayyan (Ouzzani et al., 2016). We imported all results from our literature search into Rayyan and started by removing all duplicated studies using the Rayyan smart group function for duplicates. Then, two independent reviewers (DM and MP) screened all titles, abstracts and keywords to select eligible references for further scrutiny. Any disagreements were adjudicated by consensus. The two independent reviewers (DM and MP) then assessed all potentially eligible full-texts against the inclusion/exclusion criteria to decide on the inclusion/exclusion of the reference. Any disagreements were adjudicated by consensus. The inclusion/exclusion criteria were tailored to retain all RCTs assessing the efficacy of intranasal OT, as compared with placebo, on core symptoms measured before and after a period of repeated treatment administration. The inclusion criteria were:
original article; written in English; published and peer-reviewed; placebo-controlled, double-blind, randomized clinical trial investigating the effects of repeated administration of intranasal OT on core symptoms of schizophrenia or ASD; enough available data for meta-analysis (i.e. means and SDs of pre- and post-treatment symptoms scores). We excluded: systematic reviews and meta-analyses; studies in healthy volunteers; studies not involving repeated intranasal administration of OT (i.e. single acute administrations); not placebo-controlled; not focusing on core symptoms as outcomes (e.g. using only social cognition tasks); studies where we could not retrieve sufficient data for meta-analysis.

### 2.3 | Outcomes and data extraction

Two authors independently read the full text of all included articles. We extracted the following information: first author, publication year, study design, sample size, proportion of male participants, duration of illness, diagnosis and diagnostic criteria, if patients were in- or outpatients, frequency of intranasal OT administration, daily dose, scheme of administration, total duration of treatment and the concomitant administration of any other treatment. Discrepancies were adjudicated by consensus.

For the RCTs on schizophrenia, we focused on total psychopathology (measured with the Positive and Negative Symptoms Scale (PANSS) (Kay et al., 1987), Brief Psychiatric Rating Scale (BPRS) (Fleischman & Zimmerman, 1973) or the Clinical Global Impression scale - Symptoms (CGI-S) (Busner & Targum, 2007)), or any of the three symptom sub-domains: positive symptoms (measured with BPRS, PANSS or Scale for Assessment of Positive Symptoms (SAPS) (Minas et al., 1994)), negative symptoms (measured with the PANSS, Clinical Assessment Interview for Negative Symptoms (CAINS) (Kring et al., 2013) and Scale for Assessment of Negative Symptoms (SANS) (Andreasen, 1989)) and general symptoms (measured exclusively with the PANSS). Briefly, positive symptoms include symptoms such as delusions, hallucinations and disorganized thinking; negative symptoms are characterized by deficits in cognitive, affective, and social functions, including blunting of affect and passive withdrawal; general symptoms include many deficits in cognition such as disorientation, poor attention, lack of insight and active social avoidance, that are not captured in the positive/negative symptoms domains; and total psychopathology reflects a sum of all of the above (Kay et al., 1987).

For the RCTs on ASD, we focused on global symptoms (global impression of psychopathology severity measured with the CGI-S, Childhood Autism Rating Scale (CARS) (Schopler et al., 1980) or Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 1989)), repetitive/stereotypical behaviours (measured with the Repetitive Behaviours Scale (RBS) (Lam & Aman, 2007), Autism Behaviour Checklist (ABC)-Stereotypical (Witwer & Lecavaller, 2007) or ADOS-repetitive) and social functioning (measured with the Social Responsiveness Scale (SRS) (Bolte et al., 2008), ADOS-reciprocity, Interaction Rating Scale Advanced (IRSA) (Anme et al., 2011) or scoring of social skills based on video recording of social interactions).

For all of our outcome measures, we extracted the mean and respective SDs from the pre- and post-treatment scores at each treatment group separately for the efficacy meta-analyses. If means and SDs of the changes from pre- to post-treatment were also reported, we also extracted those values for the variance meta-analyses. In crossover studies, we consider only data from the first randomization period to avoid potential treatment-by-sequence interactions. For those studies that did not report statistics on the scores changes from pre- to post-treatment, we calculated the mean difference by subtracting the post- from the pretreatment scores and the respective SD of the difference using formula 1 below (Cumpston et al., 2019):

$$SD_{change} = \sqrt{SD_{pre}^2 + SD_{post}^2 - 2SD_{pre} \times SD_{post} \times r_{pre,post}}$$  \hspace{1cm} (1)

For the schizophrenia RCTs, we assumed $r_{pre,post}$ to be 0.50 in order to be similar to correlations typically obtained in schizophrenia studies (e.g. see the Schizo_PANSS data in the R package Surrogate). For the ASD RCTs, we assumed $r_{pre,post}$ to be 0.70, as suggested by available patient data from two of the RCTs included in this study (Bernaerts et al., 2020; Muneshue et al., 2016). For completeness, to exclude that our findings might be driven by the specific assumptions we made regarding the correlation of pre- and post-treatment scores ($r_{pre,post}$), in both cases we recalculated the estimated $SD_{change}$ using three different $r_{pre,post}$ values: 0.5, 0.7 and 0.9 and repeated all of our main analyses to check whether our pooled estimates would change considerably as a function of $r_{pre,post}$. Because we did not find any considerable change in the main conclusions of our analyses, for conciseness we report only our main analyses for $r_{pre,post} = 0.50$ in schizophrenia and $r_{pre,post} = 0.70$ in ASD.

Where information was missing, we contacted the authors directly to access the missing data. If after contacting the authors, we could not gather all the information we needed to calculate the effect sizes for meta-analysis, but this information could be retrieved from available plots in the manuscript, then we used the DigitizeIt software (http://www.digitizeit.de/) to extract data presented graphically (this happened with the pretreatment social functioning scores from Parker et al., 2017). If we could not retrieve the missing data for any outcome measures from the authors or from available plots, then these outcome measures were excluded from further analyses (this happened with the pretreatment scores of repetitive/stereotypical behaviours from Parker et al., 2017).

### 2.4 | Risk of bias

We used the revised Cochrane Risk of Bias tool (RoB 2) (Higgins et al., 2011) to assess and classify the risk of bias in each of the included studies, according to criteria defined a priori. We judged whether each study had a high, low or unclear risk of bias in each of the following five domains: randomization process, deviations from intended intervention, missing outcome data, measurement of the outcome, selection of the reported results. The overall risk of bias was...
classified as low if none of the above domains were rated as high risk and three or less are rated as unclear risk. It was classified as moderate if one domain was rated as high risk, or none rated as high risk but four or more rated as unclear risk. All other studies were classified as having a high risk of bias. The risk of bias was assessed independently by two reviewers and any disparity was resolved by consensus. Risk of bias assessments for each study can be found in Figures S3 and S4.

3 | STATISTICAL ANALYSIS

3.1 | Calculation of effect sizes and respective variance

3.1.1 | Meta-analysis of efficacy

To estimate the effects of repeated intranasal OT on symptoms we used the standardized mean change using raw score standardization (SMCR) statistic (Morris, 2000). The SMCR, unlike other effect sizes such as Hedge’s g, not only contrasts treatment and control groups, but also controls for possible differences in the pretreatment values. This statistic is estimated by:

i. computing the standardized mean difference between pre- and post-treatment time points separately for the treatment and control groups, using formula 2:

\[
SMCR = J \times \frac{\bar{x}_{G \cdot OT} - \bar{x}_{G \cdot IO}}{SD_{G \cdot IO}}
\]

(2)

The indices t0 and t1 indicate different measurement times (e.g. pre- and post-treatment). The correction factor to achieve an unbiased estimator is defined as:

\[
J = 1 - \frac{3}{4(n_0 - 1) - 1}
\]

(3)

ii. the final effect size of the treatment effect is computed by calculating the difference in the SMCR between the treatment and placebo groups, using the formula below 4:

\[
SMCR = SMCR_{OT} - SMCR_{PL}
\]

(4)

Negative SMCRs values mean improvement in symptoms from pre- to post-treatment in the repeated intranasal OT group compared with placebo.

The SMCR provides a better estimate of treatment effects, but it requires knowledge of the correlations of outcomes across time points \( r_{pre\cdot post} \) to compute its variance. As described above, we assumed 0.5 for schizophrenia and 0.7 for ASD studies, but repeated all our analyses for \( r_{pre\cdot post} \) of 0.5, 0.7 and 0.9.

We estimated the variance of SMCR for each experimental group (treatment, placebo) using the formula below 5:

\[
Var(SMCR) = \frac{2(1-r_{pre\cdot post})}{n_0} + \frac{SMCR^2}{2n_0}
\]

(5)

The variance of the final treatment effect estimate is given by formula 6:

\[
Var(SMCR) = Var(SMCR_{OT}) + Var(SMCR_{PL})
\]

(6)

3.1.2 | Meta-ANOVA

For our meta-analyses of variance, we could have used either the logarithm of the variation ratio (CVR) or the coefficient of variation ratio (CVR) to estimate the relative variability in the intranasal OT and placebo groups (Senior et al., 2020). The main difference between these two metrics lies on the fact that CVR adjusts the VR for mean differences between groups, because in biological systems variance often scales with the mean; in this case, unadjusted variation ratios cannot provide an unbiased measure of variance (Senior et al., 2020). To inform the selection of the most appropriate variance effect size in our case, we examined the correlations between mean pre- to post-treatment scores changes and the respective SDs, for each treatment group and symptom domain separately. The correlation coefficients were weighted by study size using the ‘weights’ package from R (version 1.0). We used bootstrapping (1,000 samples) to assess significance. As expected, we found that in the studies we included in our meta-analyses, mean changes in symptom scores from pre- to post-treatment correlated positively with the SD of the changes in many instances (Table S1). Hence, we focused our variability meta-analyses on the CVR, using the formula below 7:

\[
CVR = \ln \left( \frac{SD_{OT}/x_{OT}}{SD_{PL}/x_{PL}} \right) + \frac{1}{2(n_0 - 1)} - \frac{1}{2(n_0 - 1)}
\]

(7)

We estimated the variance of the CVR effect size using the formula below 8:

\[
Var_{CVR} = \frac{SD^2_{OT}}{n_0 x_{OT}^2} + \frac{1}{2(n_0 - 1)} + \frac{SD^2_{PL}}{n_0 x_{PL}^2} + \frac{1}{2(n_0 - 1)} - 2\rho \sqrt{\frac{SD^2_{OT}}{n_0 x_{OT}^2} \frac{1}{2(n_0 - 1)}} - 2\rho \sqrt{\frac{SD^2_{PL}}{n_0 x_{PL}^2} \frac{1}{2(n_0 - 1)}}
\]

(8)

where \( x_{OT} \) and \( x_{PL} \) are the mean change in symptom scores, \( SD_{OT} \) and \( SD_{PL} \) the reported SDs of the difference, \( n_0 \) the sample sizes for OT and placebo treated groups respectively, and \( \rho \) the correlation between converted (log-transformed) means and SDs of changes in scores. The use of CVR to quantify group differences in variability is possible only where data have a true zero point, that is no scores lower than 0 exist. This is not the case for raw change scores which can be positive or negative. Hence, following the approach suggested...
by McCutcheon et al. (2019), we converted values of mean change to a ratio scale (using formula 9).

\[ \hat{x}_{\text{converted}} = |x_{\text{reported}}| + (x_{\text{baseline}} - C) \]  

(9)

where \( \hat{x}_{\text{converted}} \) is the converted mean change in symptoms, \(|x_{\text{reported}}|\) is the absolute value of the converted mean change in symptoms, and \(C\) represents the minimum score possible on the rating scale (e.g. 30 in the case of the PANSS). In all instances where \( \hat{x} \) is mentioned in the text this refers to \( \hat{x}_{\text{converted}} \) unless otherwise stated.

A CVR above 1 indicates greater variability in the intranasal OT arm, whereas a value below 1 indicates greater variability in the placebo arm.

3.1.3 | Meta-analysis

Our meta-analyses were carried out using the metafor package (version 2.0.0) from R. For the analyses on schizophrenia studies we used four separate univariate random effects models, one for each symptom category to obtain pooled estimates of the overall effect of repeated intranasal OT, as compared with placebo, across studies for (i) total psychopathology; (ii) positive symptoms; (iii) negative symptoms; and (iv) general symptoms. This approach was appropriate for the schizophrenia studies, given that we did not include more than one effect size from the same study in the same model, and, thus, all effect sizes in a model were independent. However, in the case of the ASD studies, one study tested two different doses. The effect sizes of these two doses cannot be considered independent because they originated from the same study. Hence, as a way of dealing with this dependence, for the ASD meta-analyses we used a hierarchical multivariate random effects model, where we specified study as a clustering factor. Applying multivariate meta-analytic models can be challenging when the covariance structure is unknown and cannot be estimated based on previous literature—which was our case. To overcome this issue, we estimated the variance–covariance matrix from the data using the clubSandwich package from R (0.5.0). We used these models to obtain separate pooled estimates of the effects of intranasal OT, as compared with placebo, on (i) global symptoms; (ii) repetitive/stereotypical behaviours; and (iii) social functioning.

For all of our meta-analytic models, we assessed the presence of influential effect sizes by calculating the Cook’s distances (we took a conservative approach where we considered influential any effect size with a Cook’s distance >0.5). When an effect size was deemed influential, we then conducted further sensitivity analyses repeating our main analysis excluding the influential effect sizes.

3.1.4 | Heterogeneity

Because studies differ more or less in their experimental design, tools used for assessing outcomes, and treatment schemes, it is likely that some heterogeneity exists between the outcomes of different studies (Xu et al., 2008). Statistical heterogeneity occurs when the true effects of the different studies show a larger variation than it should be expected due to random error or by chance. Assessing heterogeneity is an important part of interpreting the results or appraising the conclusions that can be drawn from a meta-analysis. In our study, we assessed between-studies heterogeneity using the Cochran’s Q test (Kulinskaya & Dollinger, 2015).

3.1.5 | Moderators

To evaluate the potential influence of moderators, we used meta-regression models, applied separately for each moderator. We focused on the following factors: year of publication, percentage of male participants included, mean age, design of the study (i.e. within versus between subjects design), daily dose of OT administered, daily frequency of administration, duration of treatment and the presence of a concomitant additional intervention (i.e. psychological intervention). We applied Bonferroni correction for the number of moderators investigated.

3.1.6 | Publication bias

We assessed the presence of publication bias through the inspection of funnel plots and the use of the rank correlation test for funnel plot asymmetry (Song et al., 2002).

3.1.7 | Power calculations

To inform the appraisal of our findings, we also conducted some simulations to calculate the statistical power of our different meta-analytic models to detect small (0.3), medium (0.5) and large (0.7) effect sizes under low (\( \tau^2 = 0.33 \)), moderate (\( \tau^2 = 1 \)) and large (\( \tau^2 = 3 \)) heterogeneity scenarios. We performed these simulations taking into account both the number of effect sizes and the average sample size of the studies included in each of our meta-analyses using the R script available in https://bit.ly/30lSple. These analyses would provide us with an idea of the magnitude of effect sizes our meta-analyses would be sufficiently powered to detect, given the amount of effect sizes, average sample size of the included studies and between-studies heterogeneity.

3.1.8 | Individual patient data

Because we could retrieve the raw data from two ASD RCTs (Bernaerts et al., 2020; Munesue et al., 2016), in a secondary analysis of these data we examined whether there was any evidence of deviation from unimodality in the distribution of pre- to post-treatment change scores in the repeated intranasal OT and placebo groups of
those two studies. We assessed deviations from unimodality using the Hartigan’s Dip Test of Unimodality (Kang & Noh, 2019), as implemented in the diptest package from R (0.75-7).

4 | RESULTS

4.1 | Study characteristics

4.1.1 | Schizophrenia

We found 10 eligible schizophrenia RCTs. The total number of patients randomized to receive intranasal OT and placebo were 175 and 165, respectively. Patients had mean age of 37.14 years (SD 7.38, range: 21.92–48.00) and were mostly men (mean percentage of male participants 80.98%, SD 8.58, range: 69.23–100). The daily dose of intranasal OT administered varied between 40 to 80 IU/day (this corresponded to either single administrations of 40 or administrations of 40 IU twice a day). However, most studies administered daily doses of 48 IU (24 IU twice a day). Treatment had a mean duration of 6.93 weeks (SD 4.66; range: 2–17.43). Two studies tested the effects of intranasal OT combined with concomitant social training. For an overview of the characteristics of these studies, please see Table 1.

4.1.2 | ASD

We found 10 eligible ASD RCTs. The total number of patients randomized to receive intranasal OT and placebo were 221 and 218, respectively. Patients had mean age of 20.37 years (SD 9.69, range: 6.20–33.20) and were mostly men (mean percentage of male patients 93.40%, SD 8.78, range: 78.33–100). The daily dose of intranasal OT administered varied between 12 to 48 IU/day (this corresponded to either single administrations of single doses of 12 IU or administrations of 24 IU twice a day). Most studies administered daily doses of 48 IU (24 IU twice a day). Treatment had a mean duration of 5.96 weeks (SD 3.03; range: 0.57–12). For an overview of the characteristics of these studies, please see Table 2.

4.2 | Meta-analyses

Below, we present the results of our quantitative syntheses. For clarity, we will present first the results of our meta-analyses of efficacy. Within each disorder, we will present the results for each symptoms domain in separate subsections. We will then present the results of our meta-analyses of variance, using a similar structure.

4.3 | Meta-analyses of efficacy

4.3.1 | Schizophrenia

Total psychopathology

Compared with placebo, repeated intranasal OT did not significantly improve or exacerbate total psychopathology in patients with...
schizophrenia (pooled effect size = −0.04, 95% CI [−0.31, 0.22], p = 0.75). We did not find significant heterogeneity (Q(9) = 10.77, p = 0.29). Modabbernia et al. (2013) was deemed an influential study (Cook’s distance >0.5). Repeating our meta-analysis without this study did not change the main pattern of the results (pooled effect size = 0.067, 95% CI [−0.18, 0.31], p = 0.592) (Figure 1a). Daily administered dose was the only significant moderator (β = −0.02, p = 0.009), though this did not survive Bonferroni correction (Table S2). The moderation was driven by a significant medium improvement in symptoms in the studies using 80 IU (pooled effect size = 0.61, 95% CI [1.14, 0.002], p = 0.05, k = 2 studies) whereas the 40 (pooled effect size = 0.37, 95% CI [−0.19, 0.93], p = 0.20, k = 2 studies) and 48 IU doses (pooled effect size = 0.04, 95% CI [−0.24, 0.36], p = 0.54, k = 6 studies) did not significantly improve or exacerbate symptoms. We could not find evidence of publication bias (Kendall’s τ = −0.29, p = 0.29) (Figure S5A).

Positive symptoms

Compared with placebo, repeated intranasal OT did not significantly improve or exacerbate positive symptoms in patients with schizophrenia (pooled effect size = −0.07, 95% CI [−0.33, 0.20], p = 0.62). We did not find significant heterogeneity (I² = 13.36%, Q(7) = 7.58, p = 0.37). Modabbernia et al. (2013) was deemed an influential study (Cook’s distance >0.5) (Figure 1b). Repeating our meta-analysis without this study did not change the main pattern of the results (pooled effect size = 0.05, 95% CI [−0.21, 0.31], p = 0.71). Daily administered dose was the only significant moderator (β = −0.2, p = 0.03), though this did not survive Bonferroni correction (Table S3). The moderation was driven by a significant medium improvement in symptoms in the studies using 80 IU (pooled effect size = −0.58, 95% CI [−1.17, 0.002], p = 0.05, k = 2 studies) whereas the 48 IU dose (pooled effect size = 0.09, 95% CI [−0.19, 0.36], p = 0.54, k = 6 studies) did not significantly improve or exacerbate symptoms. We could not find evidence of publication bias (Kendall’s τ = −0.29, p = 0.39) (Figure S5B).

Negative symptoms

Compared with placebo, repeated intranasal OT did not significantly improve or exacerbate negative symptoms in patients with schizophrenia (pooled effect size = −0.06, 95% CI [−0.28, 0.16], p = 0.58) (Figure 1c). We did not find significant heterogeneity (Q(9) = 2.60, p = 0.98). No study was deemed an influential study (all Cook’s distances <0.5). We did not find any significant moderator (Table S4). We could not find evidence of publication bias (Kendall’s τ = −0.20, p = 0.48) (Figure S5C).

General symptoms

Compared with placebo, repeated intranasal OT did not significantly improve or exacerbate general symptoms in patients with schizophrenia (pooled effect size = −0.06, 95% CI [−0.28, 0.16], p = 0.58) (Figure 1c). We did not find significant heterogeneity (Q(9) = 2.60, p = 0.98). No study was deemed an influential study (all Cook’s distances <0.5). We did not find any significant moderator (Table S4). We could not find evidence of publication bias (Kendall’s τ = −0.20, p = 0.48) (Figure S5C).
### TABLE 2  Characteristics of the randomized clinical trials of intranasal OT in autism spectrum disorders included in this analysis

| Study                          | N (OT) | N (PL) | Year of publication | Mean age | Percentage of men (%) | Daily dose (IU) | Scheme of administration (daily frequency) |
|-------------------------------|--------|--------|---------------------|----------|-----------------------|-----------------|------------------------------------------|
| (Anagnostou et al., 2012)     | 10     | 9      | 2012                | 33.2     | 84.21                 | 48              | Twice                                    |
| (Guastella et al., 2015)      | 26     | 24     | 2015                | 13.93    | 100.00                | 48 or 36 depending on age | Twice                                    |
| (Parker et al., 2017)         | 14     | 18     | 2017                | 8.74     | 84.38                 | 48              | Twice                                    |
| (Munesue et al., 2016)        | 15     | 14     | 2016                | 22.5     | 100.00                | 16              | Twice                                    |
| (Yamasue et al., 2020)        | 51     | 52     | 2017                | 26.95    | 100.00                | 48              | Twice                                    |
| (Watanabe et al., 2015)       | 9      | 9      | 2015                | 32.2     | 100.00                | 48              | Twice                                    |
| (Yatawara et al., 2016)       | 15     | 16     | 2016                | 6.2      | 87.1                  | 48              | Twice                                    |
| (Dadds et al., 2014)          | 19     | 19     | 2014                | 11.23    | 100.00                | 24 or 12 depending on weight (threshold 40 kg) | Single  |
| (Kosaka et al., 2016)         | 40     | 40     | 2016                | 24.27    | 78.33                 | 16 and 32       | Single                                   |
| (Bernaerts et al., 2020)      | 22     | 17     | 2020                | 24.5     | 100.00                | 24              | Single                                   |

**Note:** Please note that in the case of the crossover studies, we used only data from the first randomization period.

**Abbreviations:** ADI, R, Autism Diagnostic Interview, Revised; ADOS, Autism Diagnostic Observation Schedule; ASSQ, Autism spectrum Screening Questionnaire; BTW, between-subjects; CARS, Childhood Autism Rating Scale; CRO, crossover; DB, double blind; DISCAP, Diagnostic Interview Schedule for Children, Adolescents and Parents; DISCO, Diagnostic Interview for Social and Communication Disorders; IU, international units; OT, oxytocin; PL, placebo.

### TABLE 2 (Continued)

| Study                          | Duration of treatment (weeks) | Other psychological intervention | Blinding | Design | Diagnostic criteria | Other pharmacological treatment |
|-------------------------------|-------------------------------|---------------------------------|----------|--------|---------------------|---------------------------------|
| (Anagnostou et al., 2012)     | 6                             | No                              | DB       | BTW    | ADOS + ADI-R        | NA                              |
| (Guastella et al., 2015)      | 8                             | No                              | DB       | BTW    | ADOS + ADI-R        | 36%                             |
| (Parker et al., 2017)         | 4                             | No                              | DB       | BTW    | ADOS + ADI-R        | NA - some patients under stable treatment |
| (Munesue et al., 2016)        | 8                             | No                              | DB       | CRO    | DISCO - 11          | 22%                             |
| (Yamasue et al., 2020)        | 6                             | No                              | DB       | BTW    | ADOS + ADI-R        | 11.3%                           |
| (Watanabe et al., 2015)       | 6                             | No                              | DB       | CRO    | ADOS + ADI-R        | 0.05%                           |
| (Yatawara et al., 2016)       | 5                             | No                              | DB       | CRO    | DSM-IV-TR + ADOS    | NA - some patients under stable treatment |
| (Dadds et al., 2014)          | 0.57                          | No                              | DB       | BTW    | DISCAP-ASD + CARS + ASSQ + OSU autism rating scale: DSM IV + OSU autism global impression scale | 44.74%                           |
| (Kosaka et al., 2016)         | 12                            | No                              | DB       | BTW    | DISCO               | 31.67%                          |
| (Bernaerts et al., 2020)      | 4                             | No                              | DB       | BTW    | DSM-IV-TR + ADOS    | 20%                            |

**Note:** Please note that in the case of the crossover studies, we used only data from the first randomization period.

**Abbreviations:** ADI, R, Autism Diagnostic Interview, Revised; ADOS, Autism Diagnostic Observation Schedule; ASSQ, Autism spectrum Screening Questionnaire; BTW, between-subjects; CARS, Childhood Autism Rating Scale; CRO, crossover; DB, double blind; DISCAP, Diagnostic Interview Schedule for Children, Adolescents and Parents; DISCO, Diagnostic Interview for Social and Communication Disorders; IU, international units; OT, oxytocin; PL, placebo.
We did not find significant heterogeneity \((Q(5) = 6.00, p = 0.31)\) (Figure 1c). Both percentage of male patients included and total duration of treatment were significant moderators, though none would have survived Bonferroni correction (Table S5). Studies including higher percentages of male patients reported stronger improvements after repeated intranasal OT \((\beta = -0.10, p = 0.02)\) (Figure 2a). Studies with longer durations of treatment administration reported weaker improvements after repeated intranasal OT \((\beta = 0.08, p = 0.05)\) (Figure 2b). We could not find evidence of publication bias (Kendall’s \(\tau = -0.20, p = 0.72\) (Figure S5D)).
4.3.2 | Autism spectrum disorders

**Global symptoms**
Compared with placebo, repeated intranasal OT did not significantly improve or exacerbate global symptoms in patients with ASD (pooled effect size \(= -0.02, 95\% \text{ CI } [-0.24, 0.20], p = 0.84\) (Figure 3a). We did not find significant heterogeneity (Q(5) = 0.99, \(p = 0.96\)). No effect size was deemed influential (all Cook’s distances <0.5). We did not find any significant moderator (Table S6). We could not find evidence of publication bias (Kendall’s \(\tau = 0.07, p = 1.00\)) (Figure S6A).

**Repetitive/stereotypical behaviours**
Compared with placebo, repeated intranasal OT did not significantly improve or exacerbate repetitive/stereotypical in patients with ASD (pooled effect size = 0.03, 95\% CI \([-0.14, 0.20]\), \(p = 0.73\)). We did not find significant heterogeneity (Q(9) = 6.37, \(p = 0.73\)). The effect size from Dadds et al. (2014) was deemed influential (Cook’s distance >0.5). Excluding this study did not change the overall patterns of results (pooled effect size = 0.96, 95\% CI \([0.80, 1.16]\), \(p = 0.70\)) (Figure 3b). We did not find any significant moderator (Table S7). We could not find evidence of publication bias (Kendall’s \(\tau = -0.02, p = 1.00\)) (Figure S6B).

**Social functioning**
Compared with placebo, repeated intranasal OT did not significantly improve or exacerbate social functioning in patients with ASD (pooled effect size = -0.03, 95\% CI \([-0.14, 0.09]\), \(p = 0.065\)) (Figure 3c). We did not find significant heterogeneity (Q(9) = 16.86, \(p = 0.051\)). No effect size was deemed influential (all Cook’s distances <0.5). We did not find any significant moderator (Table S8). We could not find evidence of publication bias (Kendall’s \(\tau = 0.07, p = 1.00\)) (Figure S6C).

4.4 | Meta-analyses of variance

4.4.1 | Schizophrenia

**Total psychopathology**
We did not find any significant differences in variability between the repeated intranasal OT and placebo groups (pooled effect size = 1.18, 95\% CI \([0.92, 1.53]\), \(p = 0.19\)). We found significant heterogeneity (Q(9) = 24.92, \(p = 0.003\)). Modabbernia et al. (2013) was deemed an influential study (Cook’s distance >0.5). Repeating our meta-analysis without this study did not change the main pattern of the results (pooled effect size = 1.08, 95\% CI \([0.93, 1.27]\), \(p = 0.30\)) (Figure 4a), but we no longer found significant heterogeneity (Q(8) = 4.729.

**FIGURE 3** Efficacy of repeated intranasal oxytocin (OT) for the core symptoms of patients with autism spectrum disorders. Forest plots of the standardized mean change using raw score standardization (SMCRs) effect sizes separated by symptom’s domain. Negative effect sizes favour intranasal OT whereas positive effect sizes favour placebo. Note that here we show forest plots and summary effect sizes obtained after excluding one influential study identified in the repetitive/stereotypical behaviours set of studies (Cook’s distance >0.5). CI, confidence interval; ES, effect size.
We did not find any significant differences in variability between the repeated intranasal OT and placebo groups (pooled effect size = 1.13, 95% CI [0.67, 1.89], p = 0.09). We did not find any significant moderator. We could not find evidence of publication bias (Kendall's $\tau = -0.07, p = 0.90$) (Figure S7B).

**General symptoms**

We did not find any significant differences in variability between the repeated intranasal OT and placebo groups (pooled effect size = 0.87, 95% CI [0.74, 1.02], p = 0.09). We did not find any significant moderator (Table S12). We could not find any significant moderator. We could not find evidence of publication bias (Kendall's $\tau = -0.33, p = 0.47$) (Figure S7D).

### 4.4.2 Autism Spectrum disorders

#### Global symptoms

We did not find any significant differences in variability between the repeated intranasal OT and placebo groups (pooled effect size = 0.87, 95% CI [0.74, 1.02], p = 0.09). We did not find any significant moderator (Table S12). We could not find any significant moderator. We could not find any significant moderator. We could not find evidence of publication bias (Kendall's $\tau = -0.33, p = 0.47$) (Figure S7D).
pattern of the results (pooled effect size = 0.96, 95% CI [0.79, 1.15], p = 0.64) (Figure 5a). We did not find any significant moderator (Table S13). We could not find evidence of publication bias (Kendall’s τ = −0.20, p = 0.72) (Figure S8A).

Repetitive behaviours
We did not find any significant differences in variability between the repeated intranasal OT and placebo groups (pooled effect size = 0.88, 95% CI [0.69, 1.12], p = 0.09). We found significant heterogeneity (Q(9) = 54.95, p < 0.001). Dadds et al. (2014) was deemed an influential study (Cook’s distance >0.5). When we repeated our meta-analysis without this study, we found a significant decrease in variability in the intranasal OT as compared with the placebo group (pooled effect size = 0.81, 95% CI [0.68, 0.96], p = 0.02) (Figure 5b), but we could still find significant residual heterogeneity (Q(8) = 26.67, p < 0.001). Duration of treatment administration was a significant moderator (β = −0.09, p = 0.001), though this did not survive Bonferroni correction (Table S14). This suggests that studies with longer durations of treatment administration reported significantly less variability after intranasal OT, as compared with placebo. However, accounting for this moderator still left some unexplained heterogeneity (Q(8) = 10.07, p = 0.001). We could not find evidence of publication bias (Kendall’s τ = −0.11, p = 0.73) (Figure S8B).

Social functioning
We did not find any significant differences in variability between the repeated intranasal OT and placebo groups (pooled effect size = 0.96, 95% CI [0.88, 1.06], p = 0.42) (Figure 5c). We did not find significant heterogeneity (Q[10] = 6.47, p = 0.77). No effect size was deemed influential (all Cook’s distances < 0.5). We did not find any significant moderator (Table S15). We could not find evidence of publication bias (Kendall’s τ = 0.02, p = 1.00) (Figure S8C).

Risk of bias
The overall risk of bias of our meta-analyses was low for both schizophrenia and ASD (Figures S3 and S4).

Individual ASD patient data
Using data from Munesue et al. (2016), we found evidence of non-unimodality only for global symptoms (as measured using CARS) in the repeated intranasal OT, but not, placebo group. However, this was driven by one single patient. For repetitive/stereotypical behaviours and social functioning we did not find evidence of deviation from unimodality either in the repeated intranasal OT or placebo groups (Figure S9). Using data from Bernaerts et al. (2020), we did not find evidence of deviation from unimodality, whether in the repeated intranasal OT or placebo groups, for either repetitive/stereotypical behaviours or social functioning (Figure S10).
**Statistical power simulations**

Our power simulations demonstrated that at best our meta-analyses would have been sufficiently powered to detect medium to large, but not small, pooled effect sizes, given the amount of data we included and the overall heterogeneity panorama of our samples (Tables S16 and S17).

**5 | DISCUSSION AND CONCLUSIONS**

We conducted two sets of meta-analyses of all RCTs investigating the effects of a daily-repeated intranasal OT treatment regime on the core symptoms of patients with schizophrenia and ASD, using gold-standard methods to investigate efficacy, and heterogeneity in the response to treatment. Using the largest pool of data to date, we show that repeated treatment with intranasal OT may have a small effect in improving general symptoms of patients with schizophrenia, but has no effects on any other symptom category in schizophrenia or ASD. Additionally, we report evidence of heterogeneity in the response to repeated treatment with intranasal OT for negative symptoms in schizophrenia, suggesting that, despite lack of evidence of treatment effects at the group level, beneficial effects might still occur in a subset of responsive patients. However, we did not find evidence of heterogeneity in the response to repeated treatment with intranasal OT for other core symptoms of schizophrenia, or any of the core symptom dimensions in ASD (beyond a tentative decrease in variance after repeated intranasal OT, as compared with placebo, for repetitive/stereotypical behaviours). We discuss each of our main findings below.

Apart from general symptoms in schizophrenia, daily-repeated treatment with intranasal OT did not ameliorate or exacerbate any of the other symptoms dimensions of schizophrenia and ASD. For general symptoms in schizophrenia, we found a small, but significant, positive effect of intranasal OT in reducing symptoms, as compared with placebo, after excluding one influential study. We note that general symptoms, as assessed by the PANSS, include many aspects of human behaviour where effects of intranasal OT have generally gathered supportive evidence in a number of studies. These are, for instance, anxiety, preoccupation, tension or active social avoidance (Kay et al., 1987). Indeed, the anxiolytic effects of OT are among the most well replicated effects in both humans (Neumann & Slattery, 2016) and rodents (Yoshida et al., 2009). Although our meta-analysis did not investigate effects on specific symptoms (this data is not typically reported), our findings of a beneficial effect of intranasal OT in reducing general symptoms in patients with schizophrenia is globally consistent with the well-established role of OT in regulating neurocognitive processes potentially involved in the genesis of these symptoms, i.e. regulation of anxiety and stress response (Neumann & Slattery, 2016). Based on our findings, we suggest that future larger intranasal OT trials in schizophrenia should pay particular attention to general symptoms.

Most previous meta-analyses have concluded that repeated treatment with intranasal OT does not result in improvement in any of the main symptom categories of patients with schizophrenia or ASD (Ooi et al., 2017; Wang et al., 2019; Williams & Burkner, 2017; Zheng et al., 2019). Excluding perhaps general symptoms in schizophrenia, our findings are broadly compatible with this conclusion. The only discrepancy is reported in a recent meta-analysis that reported that intranasal OT improved total psychopathology in patients with schizophrenia, and repetitive behaviours and social functioning (but not global symptoms) in patients with ASD (Peled-Avron et al., 2020). This discrepancy most likely reflects the fact that the authors used effect sizes that did not control for pretreatment symptom scores. Hence, our study provides the most updated and unbiased quantitative synthesis of the evidence on the effects of repeated treatment with intranasal OT on the core symptoms of patients with schizophrenia and ASD.

Our analyses of moderators highlighted a number of factors that might explain some of the differences across studies regarding the effect of repeated treatment with intranasal OT. In alignment with Zheng et al. (2019), we also found that dose significantly moderated the effects of intranasal OT on total psychopathology and positive symptoms in schizophrenia. Here, studies using the highest daily administered dose (80 IU) reported medium improvements in total psychopathology and positive symptoms as compared with placebo; this effect was not present in 40 and 48 IU daily doses used in other studies. However, we urge for caution in interpreting these findings because these significant moderating effects were driven by two studies using large 80 IU daily doses and one of these studies was deemed influential in both meta-analyses. These findings suggest that further work is required to elucidate dose–response curves in improving psychopathology in patients with schizophrenia before drawing conclusions with any confidence.

Interestingly, we also found that the percentage of male patients included and total duration of treatment were significant moderators of the effects of repeated treatment with intranasal OT on general symptoms of schizophrenia. Here, studies including higher percentages of male patients tended to report larger beneficial effects of intranasal OT; studies with longer periods of treatment administration tended to report smaller improvements after intranasal OT. These findings are interesting for two reasons. First, gender is a well-known moderator of the effects of intranasal OT (Lynn et al., 2014; Xu et al., 2017), including on anxiety and stress response (Steinman et al., 2016), which contribute to the genesis of general symptoms in schizophrenia. For instance, although studies in men have consistently reported anxiolytic effects of intranasal OT (Heinrichs et al., 2003; Spengler et al., 2017), studies in women have reported opposite effects (Domes et al., 2007, 2010; Lischke et al., 2012). Hence, it is plausible that intranasal OT might attenuate more strongly general symptoms in male patients, but have no effects or even exacerbate these symptoms in female patients. This would fit well with our observation that studies including higher percentages of male patients report larger improvements in general symptoms after schizophrenia. Further studies should investigate this hypothesis.

Second, there has been intense debate about the optimal scheme of administration when treatment with intranasal OT involves
repeated applications (Horta et al., 2020). Given the high propensity of the OT receptor to desensitization (Robinson et al., 2003), long-term exposure to high concentrations of OT might down-regulate the OT signalling machinery and thus attenuate the effect of OT on brain function. Down-regulation of OTR expression after chronic exposure to OT has been clearly demonstrated in rodents (Huang et al., 2014; Pisansky et al., 2017). Supporting this idea, a recent study investigating the impact of dose frequency on the effects of intranasal OT in healthy volunteers found that the repeated administration of intranasal OT every other day over five days was more effective in dampening the response of the amygdala to negative affect, when compared with daily administrations (Kou et al., 2020). This hypothesis fits well with our observation that studies with longer periods of administration reported smaller improvements in general symptoms after intranasal OT. It is imperative that future studies systematically investigate the efficacy of various dosing and frequency of administration regimes for RCTs using repeated applications of intranasal OT. Suboptimal regimes can only increase inconsistencies in reported findings, reducing our confidence in the role of intranasal OT as a potential treatment for psychopathology.

For the first time, using meta-analyses of variance to investigate whether a repeated intranasal OT treatment regime might have beneficial effects for subsets of responsive patients, we report evidence of increased variance in the intranasal OT group, compared with the placebo group, for negative symptoms in schizophrenia. This suggests the presence of heterogeneity in treatment response, which in turn might reflect the existence of a subgroup of patients that respond positively to intranasal OT with respect to improvement in negative symptoms, even though the overall mean group effect was not significant. Our meta-analysis cannot identify which factors might determine which patients will respond to intranasal OT in terms of negative symptoms. This needs to be elucidated in future RCTs, for example by using a repeated N-of-1 design, where the effects of treatment are studied by following an individual patient over time with the treatments given being randomised from period to period (Araujo et al., 2016). Parsing this heterogeneity will be important to identify treatment response biomarkers which could guide the selection of patients with higher likelihood of responding to intranasal OT. This is particularly important for negative symptoms, which currently still lack highly effective treatments (Remington et al., 2016). In the absence of this knowledge, future RCTs based on heterogeneous cohorts of patients will, most likely, result in inconclusive findings and missed opportunities to advance the treatment of negative symptoms in schizophrenia by targeting the OT system.

Although the exact factors underlying the heterogeneity in the response to intranasal OT treatment for negative symptoms in schizophrenia remain unclear, we can postulate some hypotheses here. First, negative symptoms encompass varied behavioural problems that include blunted affect, emotional withdrawal, poor rapport, apathetic social withdrawal, difficulties in abstract thinking, lack of spontaneity and stereotyped thinking (Kay et al., 1987). It is conceivable that given the preferential effects of OT on social affiliation (Feldman, 2012), intranasal OT might benefit only patients where social deficits contribute the most to the negative symptoms score, whereas those patients where executive deficits prevail might not benefit as much. Second, a number of inter-individual factors have already been demonstrated to moderate the effects of intranasal OT. Some examples include, for instance, genotype (e.g. the single nucleotide polymorphism rs53576 in the promoter of the OXTR gene) (Chen et al., 2020; Feng et al., 2015) and early life adversity (Joseph et al., 2019; van Zuiden et al., 2017), the latter being common in patients with schizophrenia (Popovic et al., 2019). Most likely these factors moderate the response to intranasal OT by affecting the regulation of the expression of OT receptors (Kraaijenvanger et al., 2019). Although the influence of genetic factors is somewhat easy to conceive, i.e. genetic polymorphisms might increase/decrease the expression of the receptor or promote/inhibit signalling without changes in expression, understanding how environmental context might shape the OT system is more challenging. Fortunately, recent evidence has shown that environmental factors might impact, for instance, on the expression of the OT receptor through epigenetic regulation (Perkeybile et al., 2019). This epigenetic regulation has been suggested as a mechanism underlying the consequences of early life adversity on the OT system (Kraaijenvanger et al., 2019; Perkeybile et al., 2019). Hence, we suggest that future larger trials start to consider the role these inter-individual factors might play in determining the clinical response to intranasal OT.

Our study did not yield evidence of increased variance in the response to intranasal OT treatment, compared with placebo, for other core symptoms of schizophrenia, or any of the core symptom dimensions in ASD (beyond a tentative decrease in variance after repeated intranasal OT, as compared with placebo, for repetitive/stereotypical behaviours). The similarity in variances in the intranasal OT and placebo groups does not support the idea that different groups of responsive and non-responsive patients might co-exist. Instead, it is more likely that variance in these cases simply reflects inter-individual differences in the course of symptom progression, which was similar for the two groups. In ASD, this conclusion was further supported by analyses on individual patient data from two RCTs, where we did not find systematic of evidence of non-unimodality in the response to intranasal OT. We note though that for repetitive/stereotypical behaviours in ASD, we found tentative evidence for decreased variance after repeated intranasal OT, as compared with placebo, suggesting that treatment with repeated intranasal OT might make repetitive/stereotypical behaviours in these patients more uniform. One possible explanation is that repeated intranasal OT might have had a stabilizing effect (Cortes et al., 2018). One example for such variance stabilization might be via ceiling or floor effects, where the assessment instrument is too coarse to capture patients’ improvements over a certain level (Cortes et al., 2018). Interestingly, duration of treatment was a significant moderator of variance in the response of repetitive/stereotypical behaviours to repeated intranasal OT, where we found that studies with longer treatment durations tended to report less variance than those with shorter durations. This observation is in keeping with the idea that treatment duration might play
an important role in moderating the response to repeated intranasal OT.

Certain limitations of our study and of existing RCTs provide caveats that need to be considered, and where possible investigated in future studies, before any firm conclusions can be drawn. First, the number of intranasal OT RCTs is still relatively small for both disorders. We also note that these studies are mostly small-scale studies, including rather small cohorts of patients. These two aspects pose important limitations to statistical power. According to our statistical power simulations, our meta-analyses were sufficiently powered to detect significant medium-to-large, but not small, effect sizes if they existed. Second, although we found no robust evidence for differences in variance between intranasal OT and placebo for most of the schizophrenia and ASD symptom domains that we examined, intervention-by-individual interactions are still possible even if the variance ratio is equal to one. This would correspond to a scenario where subsets of patients might show high variance in their response, whereas in others, treatment may lead to a more uniform response (i.e. less variance). However, at least for ASD, this is unlikely given that in our analyses of modality of individual patient data from two RCTs, we could not find systematically robust evidence of deviations from unimodality in symptoms. Third, we were not able to investigate the effects of other concomitant treatments (i.e. antipsychotics) might have on the response to OT because descriptions of medication regimens were too sparse and too varied in the primary studies. Fourth, although not specific to this study, we should highlight that, as a general limitation in the field, we remain unclear about the exact pharmacological actions driving potential changes in symptoms in patients, if they exist. There has been intense debate on whether the effects of intranasal OT modulate behaviour through direct effects on the brain, effects on peripheral organs or both (Leng & Ludwig, 2016). Supporting the central effects hypothesis, recent evidence has shown that (i) exogenous OT, when administered intranasally, can reach the brain in primates (Lee et al., 2020); (ii) intranasal OT modulates brain physiology beyond those effects produced by intravenous OT (Martins et al., 2020; Martins, Dipasquale, et al., 2021); (iii) and that the effects of intranasal OT on the physiology of the living human brain can be predicted by the distribution of the OXTR mRNA expression in the post-mortem human brain (Martins, Broadmann, et al., 2021), which provides indirect support to the idea that intranasal OT can target the brain OT system in humans. However, we have also recently shown that exogenous OT when administered intravenously can modulate brain physiology (Martins et al., 2020) and connectivity (Martins, Dipasquale, et al., 2021) with patterns that partly overlap with those changes produced by intranasal OT. The exact mechanisms driving these changes in the brain after intravenous OT remain unknown, but recent studies in rodents have shown that they might involve signalling through the vagus nerve (Everett et al., 2021) or RAGE-mediated transport across the blood-brain barrier (Yamamoto et al., 2019). The exact source of the effects of intranasal OT on the brain and behaviour will need to be dissected in future studies combining intranasal OT with selective brain-penetrant and non-brain penetrating antagonists to isolate the potential contribution of peripheral actions, if they exist.

In summary, our meta-analysis presents evidence for a limited therapeutic effect of a repeated intranasal OT treatment regime, possibly restricted to general symptoms in schizophrenia. We identify a number of important factors, such as gender, dose and frequency, pattern and duration of repeated intranasal OT administration, and the effects of these factors must be elucidated in future studies. Our study further reports that for negative symptoms in schizophrenia, despite the lack of effects at the group-level, there is evidence for heterogeneous response to treatment in patients treated with intranasal OT. This is compatible with the idea that at least some patients might respond positively to intranasal OT. Further studies dissecting this heterogeneity will be paramount in providing a better appraisal of the potential of intranasal OT in treating negative symptoms in schizophrenia.

5.1 Nomenclature of targets and ligands

Key protein targets and ligands in this article are hyperlinked to corresponding entries in the IUPHAR/BPS Guide to PHARMACOLOGY (http://www.guidetopharmacology.org), and are permanently archived in the Concise Guide to PHARMACOLOGY 2019/20 (Alexander, Christopoulos et al., 2019; Alexander, Kelly et al., 2019).

ACKNOWLEDGEMENTS

D. M. is supported by the National Institute for Health Research (NIHR), Biomedical Research Centre at South London and Maudsley NHS Foundation Trust, and King’s College London.

CONFLICT OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

ORCID

Yannis Paloyelis http://orcid.org/0000-0002-4029-3720

REFERENCES

Alexander, S. P. H., Christopoulos, A., Davenport, A. P., Kelly, E., Mathie, A., Peters, J. A., Veale, E. L., Armstrong, J. F., Faccenda, E., Harding, S. D., Pawson, A. J., Sharman, J. L., Southan, C., Davies, J. A., & CGTP Collaborators. (2019). The Concise Guide to PHARMACOLOGY 2019/20: G protein-coupled receptors. British Journal of Pharmacology, 176(Suppl 1), S21–S5141.

Alexander, S. P. H., Kelly, E., Mathie, A., Peters, J. A., Veale, E. L., Faccenda, E., Harding, S. D., Pawson, A. J., Sharman, J. L., Southan, C., Buneman, O. P., Cidlowski, J. A., Christopoulos, A., Davenport, A. P., Fabbro, D., Spedding, M., Striessnig, J., Davies, J. A., & CGTP Collaborators. (2019). The Concise Guide to PHARMACOLOGY 2019/20: Other Protein Targets. British Journal of Pharmacology, 176, S1–S20. https://doi.org/10.1111/bph.14747
randomised trials. BMJ (Clinical Research Ed), 343, d5928. https://doi.org/10.1136/bmj.d5928

Horta, M., Kaylor, K., Feifel, D., & Ebner, N. C. (2020). Chronic oxytocin administration as a tool for investigation and treatment: A cross-disciplinary systematic review. Neuroscience and Biobehavioral Reviews, 108, 1–23. https://doi.org/10.1016/j.neubiorev.2019.10.012

Huang, H., Michetti, C., Busnelli, M., Manago, F., Sannino, S., Scheggia, D., Giancardo, L., Sona, D., Murino, V., Chini, B., Scattoni, M. L., & Papaleo, F. (2014). Chronic and acute intranasal oxytocin produce divergent social effects in mice. Neuropsychopharmacology, 39, 1102–1114. https://doi.org/10.1038/npp.2013.310

Jarkos, L. F., Pedersen, C. A., Johnson, J. L., Hamer, R. M., Rau, S. W., Elliott, T., & Dell, D. L. (2017). A 12-week randomized controlled trial of twice-daily intranasal oxytocin for social cognitive deficits in people with schizophrenia. Schizophrenia Research, 185, 88–95. https://doi.org/10.1016/j.schres.2017.01.008

Joseph, J. E., Vaughan, B. K., Camp, C. C., Baker, N. L., Sherman, B. J., Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophrenia Bulletin, 13, 261–276. https://doi.org/10.1093/schbul/13.2.261

Kosaka, H., Okamoto, Y., Munesue, T., Yamase, H., Inohara, K., Fujikoa, T., Anne, T., Orihaka, M., Ishiob, M., Jung, M., Fujisawa, T. X., Tanaka, S., Arai, S., Asano, M., Saito, D. N., Sadato, N., Tomoda, A., Omori, M., Sato, M., ... Wada, Y. (2016). Oxytocin efficacy is modulated by dosage and oxytocin receptor genotype in young adults with high-functioning autism: A 24-week randomized clinical trial. Translational Psychiatry, 6, e872. https://doi.org/10.1038/tp.2016.152

Kou, J., Zhang, Y., Zhou, F., Sindermann, C., Montag, C., Becker, B., & Kendrick, K. M. (2020). A randomized trial shows dose-frequency and genotype may determine the therapeutic efficacy of intranasal oxytocin. Psychological Medicine, 1–10. https://doi.org/10.1017/S0033291720003802

Kraaijvanger, E. J., He, Y., Spencer, H., Smith, A. K., Bos, P. A., & Boks, M. P. M. (2019). Epigenetic variability in the human oxytocin receptor (OXTR) gene: A possible pathway from early life experiences to psychopathologies. Neuroscience and Biobehavioral Reviews, 96, 127–142. https://doi.org/10.1016/j.neubiorev.2018.11.016

Kring, A. M., Pedersen, C. A., Johnson, J. L., Hamer, R. M., Rau, S. W., Elliott, T., & Dell, D. L. (2017). A 12-week randomized controlled trial of twice-daily intranasal oxytocin for social cognitive deficits in people with schizophrenia. Schizophrenia Research, 185, 88–95. https://doi.org/10.1016/j.schres.2017.01.008

Kang, Y.-J., & Noh, Y. (2019). Development of Hartigan’s dip statistic with bimodality coefficient to assess multimodality of distributions. Mathematical Problems in Engineering, 2019, 4819475.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophrenia Bulletin, 13, 261–276. https://doi.org/10.1093/schbul/13.2.261

Lord, C., Rutter, M., Goode, S., Heenensbergen, J., Jordan, H., Mawhood, L., & Scholper, E. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. Journal of Autism and Developmental Disorders, 19, 185–212. https://doi.org/10.1007/BF02211841

Lynn, S. K., Hoge, E. A., Fischer, L. E., Barrett, L. F., & Simon, N. M. (2014). Gender differences in oxytocin-associated disruption of decision bias during emotion perception. Psychiatry Research, 219, 198–203. https://doi.org/10.1016/j.psychres.2014.04.031

Martins, D., Broadmann, K., Veronese, M., Dipasquale, O., Mazibuko, N., Schuschug, U., Zelaya, F., Fotopoulou, K., & Paloyelis, Y. (2021). “Less is more”: A dose-response mechanistic account of intranasal oxytocin pharmacodynamics in the human brain. bioRxiv. 2021.001.0018.427062

Martins, D., Dipasquale, O., & Paloyelis, Y. (2021). Oxytocin modulates local topography of human functional connectome in healthy men at rest. Communications Biology, 4, 68. https://doi.org/10.1038/s42003-020-01610-z

Martins, D., Mazibuko, N., Zelaya, F., Vasilakopoulou, S., Loveridge, J., Oates, A., Maltezos, S., Mehta, M., Wastling, S., Howard, M., McAlonan, G., Murphy, D., Williams, S. C. R., Fotopoulou, A., Schuschug, U., & Paloyelis, Y. (2020). Effects of route of administration on oxytocin-induced changes in regional cerebral blood flow in humans. Nature Communications, 11, 1160. https://doi.org/10.1038/s41467-020-14845-5

McCUTCHEON, R. A., PILLINGER, T., MIZUNO, Y., MONTAGH, A., PANDIAN, H., VANO, L., MARQUES, T. R., & HOWES, O. D. (2019). The efficacy and heterogeneity of antipsychotic response in schizophrenia: A meta-analysis, Molecular Psychiatry, 26, 1–11. https://doi.org/10.1038/s41380-019-0502-5

Minas, I. H., Klimidis, S., Stuart, G. W., Copolov, D. L., & Singh, B. S. (1994). Positive and negative symptoms in the psychoses: Principal components analysis of items from the scale for the assessment of positive symptoms and the scale for the assessment of negative symptoms. Comprehensive Psychiatry, 35, 135–144. https://doi.org/10.1016/0010-440X(94)90059-Q

Mizuno, Y., McCUTCHEON, R. A., BRUGGER, S. P., & HOWES, O. D. (2020). Heterogeneity and efficacy of antipsychotic treatment for schizophrenia with or without treatment resistance: A meta-analysis. Neuropsychopharmacology, 45, 622–631. https://doi.org/10.1038/s41386-019-0577-3

Modabbernia, A., Rezaei, F., Salehi, B., Jafarinia, M., Ashrafi, M., Tabrizi, M., Hosseini, S. M. R., Tajdini, M., Ghaleiha, A., & Akhondzadeh, S. (2013). Intranasal oxytocin as an adjunct to risperidone in patients with schizophrenia: An 8-week, randomized, double-blind, placebo-controlled study. CNS Drugs, 27, 57–65. https://doi.org/10.1007/s40263-012-0022-1

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. International Journal of Surgery, 8, 336–341.
Xu, H., Platt, R. W., Luo, Z. C., Wei, S., & Fraser, W. D. (2008). Exploring heterogeneity in meta-analyses: Needs, resources and challenges. *Pediatric and Perinatal Epidemiology, 22*(Suppl 1), 18–28. https://doi.org/10.1111/j.1365-3016.2007.00908.x

Xu, X., Yao, S., Xu, L., Geng, Y., Zhao, W., Ma, X., Kou, J., Luo, R., & Kendrick, K. M. (2017). Oxytocin biases men but not women to restore social connections with individuals who socially exclude them. *Scientific Reports, 7*, 40589. https://doi.org/10.1038/srep40589

Yamamoto, Y., Liang, M., Munesue, S., Deguchi, K., Harashima, A., Furuhara, K., Yuhi, T., Zhong, J., Akther, S., Goto, H., Eguchi, Y., Kitao, Y., Hori, O., Shiraiishi, Y., Ozaki, N., Shimizu, Y., Kamide, T., Yoshikawa, A., Hayashi, Y., … Higashida, H. (2019). Vascular RAGE transports oxytocin into the brain to elicit its maternal bonding behaviour in mice. *Communications Biology, 2*, 76. https://doi.org/10.1038/s42003-019-0325-6

Yamasue, H., Okada, T., Munesue, T., Kuroda, M., Fujioka, T., Uno, Y., Matsumoto, K., Kuwabara, H., Mori, D., Okamoto, Y., Yoshimura, Y., Kawakubo, Y., Ario, Y., Kojima, M., Yuhi, T., Owada, K., Yassin, W., Kushima, I., Benner, S., … Kosaka, H. (2020). Effect of intranasal oxytocin on the core social symptoms of autism spectrum disorder: A randomized clinical trial. *Molecular Psychiatry, 25*, 1849–1858. https://doi.org/10.1038/s41380-018-0097-2

Yatawara, C. J., Einfeld, S. L., Hickie, I. B., Davenport, T. A., & Guastella, A. J. (2016). The effect of oxytocin nasal spray on social interaction deficits observed in young children with autism: A randomized clinical crossover trial. *Molecular Psychiatry, 21*, 1225–1231. https://doi.org/10.1038/mp.2015.162

Yoshida, M., Takayanagi, Y., Inoue, K., Kimura, T., Young, L. J., Onaka, T., & Nishimori, K. (2009). Evidence that oxytocin exerts anxiolytic effects via oxytocin receptor expressed in serotonergic neurons in mice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 29*, 2259–2271. https://doi.org/10.1523/JNEUROSCI.15593-08.2009

Young, L. J. (2015). Oxytocin, social cognition and psychiatry. *Neuropsychopharmacology, 40*, 243–244. https://doi.org/10.1038/npp.2014.186

Zheng, W., Zhu, X. M., Zhang, Q. E., Yang, X. H., Cai, D. B., Li, L., Li, X. B., Ng, C. H., Ungvari, G. S., Ning, Y. P., & Xiang, Y. T. (2019). Adjunctive intranasal oxytocin for schizophrenia: A meta-analysis of randomized, double-blind, placebo-controlled trials. *Schizophrenia Research, 206*, 13–20. https://doi.org/10.1016/j.schres.2018.12.007

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Martins, D., Paduraru, M., & Paloyelis, Y. (2021). Heterogeneity in response to repeated intranasal oxytocin in schizophrenia and autism spectrum disorders: A meta-analysis of variance. *British Journal of Pharmacology, 1–19*. https://doi.org/10.1111/bph.15451