ORIGINAL PAPER

Is Pervasive Developmental Disorder Not Otherwise Specified Less Stable Than Autistic Disorder? A Meta-Analysis

Emélie Rondeau · Leslie S. Klein · André Masse ·
Nicolas Bodeau · David Cohen · Jean-Marc Guilé

Abstract We reviewed the stability of the diagnosis of pervasive developmental disorder not otherwise specified (PDD-NOS). A Medline search found eight studies reiterating a diagnostic assessment for PDD-NOS. The pooled group included 322 autistic disorder (AD) and 122 PDD-NOS cases. We used percentage of individuals with same diagnose at Times 1 and 2 as response criterion. The pooled Relative Risk was 1.95 ($p < 0.001$) showing that AD diagnostic stability was higher than PDD-NOS. When diagnosed before 36 months PDD-NOS bore a 3-year stability rate of 35%. Examining the developmental trajectories showed that PDD-NOS corresponded to a group of heterogeneous pathological conditions including prodromic forms of later AD, remitted or less severe forms of AD, and developmental delays in interaction and communication.

E. Rondeau · A. Masse · J.-M. Guilé (✉)
Université de Montréal and Hôpital Rivière-des-Prairies,
7070, boulevard Perras, Montréal, QC H1E 1A4, Canada
e-mail: jmguile@total.net

L. S. Klein · J.-M. Guilé
Autism Spectrum Disorders Clinic, Child & Adolescent
Psychiatry Division, Douglas Mental Health University Institute,
Montreal, QC, Canada

N. Bodeau · D. Cohen · J.-M. Guilé
Service de psychiatrie de l'enfant et de l'adolescent,
GH Pitié-Salpêtrière, Paris, France

D. Cohen
Pierre et Marie Curie University of Paris, Paris, France

J.-M. Guilé
Child & Adolescent Psychiatry Division, Mc Gill University,
Montréal, QC, Canada

Keywords Validity · Diagnosis · Autistic disorder ·
Pervasive developmental disorder · Autism · Meta-analysis

Introduction

Over the past 15 years, there has been increasing interest in the early identification of autism spectrum disorders (ASD). In that respect, several studies have examined the stability of early diagnosis (Lord 1995; Cox et al. 1999; Moore and Goodson 2003; Charman et al. 2005). In keeping with those studies, we conducted a meta-analysis focussing on the stability of the diagnosis of pervasive developmental disorder not otherwise specified (PDD-NOS), when diagnosed for the first time in young children.

Three diagnostic categories are included in ASD: autistic disorder (AD), pervasive developmental disorder not otherwise specified (PDD-NOS) and Asperger's syndrome (AS). PDD-NOS represents an important ASD subtype due to the high frequency with which it is diagnosed, even though it is the least studied (Matson and Boisjoli 2007). In fact, prevalence estimates are 2.6 per 10,000 for AS, 13 per 10,000 for AD and 20.8 per 10,000 for PDD-NOS (Fombonne 2005). Despite PDD-NOS tending to be more commonly diagnosed than AD, there are very few studies on the predictive validity and stability of this diagnostic category.

Early identification of ASD facilitates participation in specialized intervention programs. Studies of the impact of early intervention reported significant improvements in communication skills and social behaviour, and diminished maladaptive behaviours (Charman and Baird 2002). The eligibility for participation in these programs is often limited to children with a formal diagnosis of autism, supporting the relevance of an accurate early diagnosis. However, earlier identification of ASD highlights

important issues such as the difficulty to diagnose ASD and to differentiate AD and PDD-NOS in very young children.

Diagnosing autism at a very young age is possible but clinically challenging because of its overlap with severe language delays or disorders and general developmental delays due to mental retardation (Lord 1995; Charman and Baird 2002). Van Daalen et al. (2009) reported in their study that even experienced clinicians had disagreements on the distinction between ASD and an intellectual disability without ASD in 2-year-old children. Lord (1995) also reported that clearer diagnostic differences between autistic and non-autistic children occur at age 3. At this age, over-inclusion of children with severe mental handicap and severe communication difficulties as autistic was less frequent than at age 2 (Lord 1995). This lack of predictive validity of ASD diagnoses might be due in part to developmental issues. In fact, studies of family home movies reveal that stereotypical behaviours and restricted interests appear later in the course of ASD, impeding early diagnosis (Saint-Georges et al. 2010).

Although the differentiation between delayed and deviant development remains clinically challenging in the first 2 years of life, inter-rater reliability for the distinction between ASD and non-ASD tend to be good to excellent (Van Daalen et al. 2009). Mahoney et al. (1998) found that experienced clinicians could reliably differentiate those with and without ASD ($\kappa = 0.67$). This result is consistent with the DSM-IV autism field trial (Klin et al. 2000) which reported an excellent inter-rater reliability ($\kappa = 0.95$) for the distinction between AD and non-PDD diagnoses. More recently, Van Daalen et al. (2009) found a good agreement between psychiatrists ($\kappa = 0.74$) for the distinction between ASD and non-ASD in 2-year-old children. The finer distinction between ASD subtypes, especially AD and PDD-NOS, seems more problematic with authors consistently reporting lower inter-rater agreements: $\kappa = 0.18$ (Mahoney et al. 1998); $\kappa = 0.65$ (Klin et al. 2000); $\kappa = 0.51$ (Van Daalen et al. 2009). These values are consistent with Witwer and Lecavalier (2008) who concluded that clinicians and researchers were not able to discriminate the three ASD subtypes based on the current DSM criteria. In addition to these observations on the inter-rater reliability of the PDD-NOS diagnosis, concerns have been raised about its predictive validity. While several studies have reported high stability levels for AD when diagnosed at age 2 (Lord 1995; Cox et al. 1999; Moore and Goodson 2003; Charman et al. 2005), lower stability has been observed for PDD-NOS (Cox et al. 1999). Walker et al. (2004) suggested that these results originate in the inability to accurately diagnosed PDD-NOS with current autism assessment instruments and the ambiguous DSM-IV-TR conceptualization.

In order to reinforce the reliability and validity of the diagnostic categories, a number of standardised instruments including the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R) have been developed over the past years. The ADOS is a standardised diagnostic instrument which provides a direct observation of the child in different domains: social interactions, communication, play and imagination, and stereotyped behaviours. Each domain has cut-off scores for classification, and the scores for the first two domains are totalled to provide a summary score. Children can be classified as autistic, PDD-NOS or non-autistic (Lord et al. 1999). The ADI-R is a clinician-based parent interview that evaluates the child's communication, social development, play and restricted, repetitive and stereotyped behaviours. Children can be classified as autistic or non-autistic (Lord et al. 1994). Diagnostic evaluation relies mostly on the clinical presentation around 4–5 years of age. Thus the utility of using this instrument in very young children has been questioned (Charman and Baird 2002). Another disadvantage about the use of ADI-R is that there is no scoring algorithm for PDD-NOS or AS (Scahill 2005). Despite the development of these empirically validated diagnostic instruments, clinical judgment by experienced clinicians remains the gold standard for an ASD diagnosis, especially in very young children (Charman and Baird 2002). In regards to those limitations, standard diagnostic instruments should be used with caution.

According to DSM-IV-TR (2000), PDD-NOS is to be used when there is a social interaction impairment associated with either communication impairments or with the presence of repetitive and stereotyped behaviours, interests, and activities, but the criteria for another PDD are not met. According to DSM-IV-TR criteria, PDD-NOS remains a heterogeneous group because it includes all children in the autism spectrum who are not classified with AD or AS. Thus, some concerns have been raised regarding the validity of the PDD-NOS category, with many questioning the vague criteria and the heterogeneous group of individuals included in this diagnostic category (Towbin 2005). Walker et al. (2004) also highlighted the fact that children with many diverse clinical and heterogeneous features are given the PDD-NOS diagnosis due to a lack of a better clinical definition. In regards to those concerns, Witwer and Lecavalier (2008) reviewed studies that examined differences between ASD subtypes but they were not able to conclude on the validity of the three ASD subtypes due to the inconsistent application of diagnostic criteria by researchers. Therefore, the validity of the distinction between ASD subtypes still remains a topic of debate among clinicians and researchers. Given the low inter-rater reliability for the distinction between AD and PDD-NOS, Van Daalen et al. (2009) recently questioned the validity and utility to differentiate PDD-NOS and AD at a very young age. They suggested restricting clinical diagnosis to ASD or non-ASD in 2-year-old children. This is consistent with Walker et al. (2004) who reported that the

differentiation of ASD subtypes should be done in children older than 36 months to allow for the repetitive behaviours and cognitive and language impairments to be apparent.

Thus, as PDD-NOS refers to a heterogeneous construct associated with several diagnostic validity concerns, many clinicians tend to see this diagnostic category as supposedly less stable over time (Towbin 2005). However, this assumption should be challenged with respect to the state of the literature. This paper aims at examining the temporal stability of the PDD-NOS category using a meta-analytic method, exploring the developmental trajectories of PDD-NOS over time and finally bringing light on the reliability of early PDD-NOS diagnoses in young children.

Methods

Study Inclusion

The literature search was conducted in the Ovid Medline database to locate published articles from January 1996 to October 2009. We obtained 180 articles by combining a diagnostic term (pervasive developmental disorder/diagnosis or autistic disorder/diagnosis) with a descriptor term (prospective studies or longitudinal studies or early

diagnosis or follow-up studies or outcome assessment). To be included in this meta-analysis, studies had to follow a longitudinal study design, provide original information about PDD-NOS diagnostic stability and use DSM criteria. Based on these criteria, we selected 8 articles for inclusion in the meta-analysis (see Fig. 1 article selection process). Most studies yielded by the electronic search were about AD and Asperger's Disorder rather than PDD-NOS. Consequently, we obtained a low number of articles for this review.
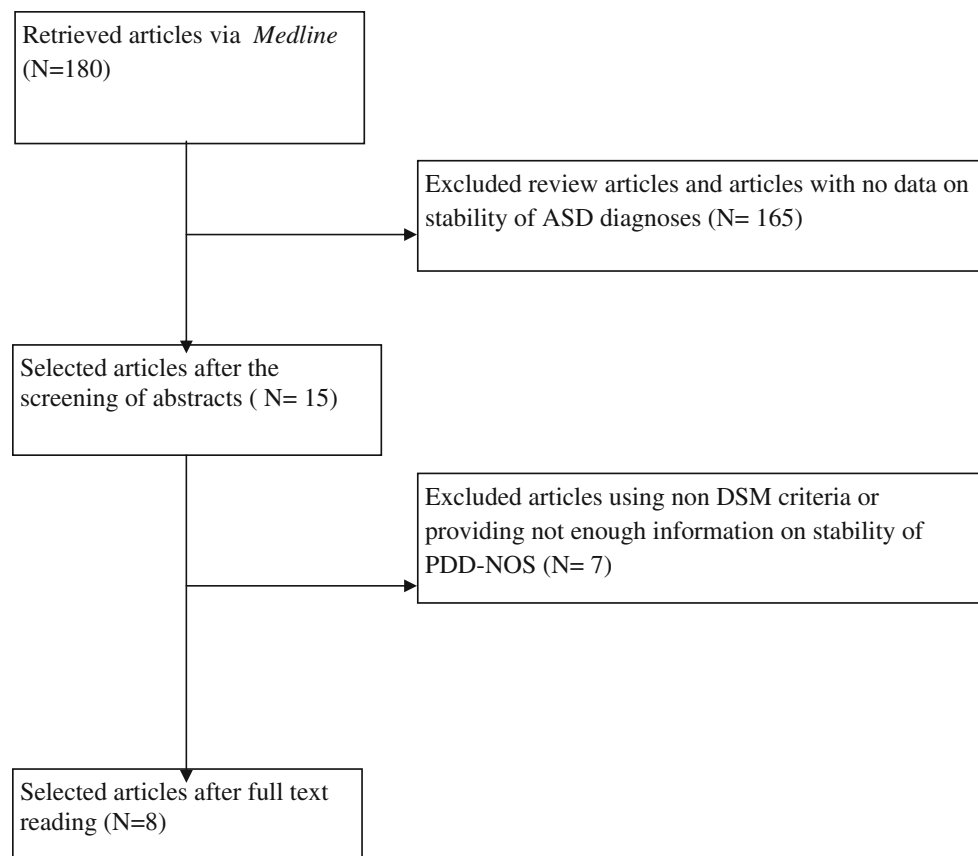
Response Criteria and Hypothesis

Exploring diagnostic stability, the response criterion chosen for the meta-analysis was the percentage of individuals with the same diagnosis at Times 1 and 2. We hypothesized that PDD-NOS was less stable than AD.

Statistical Analysis

The statistical analysis was performed with EasyMA software (Cucherat et al. 1997). We used the relative risk (RR) as a parameter of stability. A random model was chosen as the chi-square heterogeneity test first performed, yielded a value of 21.56 ($p = 0.003$).

Fig. 1 Article selection process. *AD* autistic disorder, *PDD-NOS* pervasive developmental disorder not otherwise specified, *ASD* autism spectrum disorder

Results

Demographic and Assessment Characteristics of the Studies

After combining all the studies, we obtained a total sample of 322 AD children and 122 PDD-NOS children. In all studies the first diagnostic assessment was conducted before 36 months. Mean age at Time 1 ranged from 21.6 to 33.0 months and mean age at Time 2 ranged from 34.8 to 113.8 months. Time intervals ranged from 12 to 84 months. All studies used clinical diagnoses based on DSM-IV criteria. Two of them used DSM-IV-TR diagnostic criteria (Turner et al. 2006; Turner and Stone 2007). All studies except Stone et al. (1999) and Eaves and Ho (2004) used the ADI-R and all studies but Eaves and Ho (2004) used the ADOS as diagnostic measures. Turner and Stone (2007) used the ADI-R at Time 2 only, Chawarska et al. (2007) used the ADI-R at Time 1 only and Stone et al. (1999) used ADOS at Time 2 only. Inception of treatments between Times 1 and 2 was reported in all studies.

Diagnostic Stability

As mentioned in the Method section, we used the relative risk (RR) as a parameter of stability over time. When the eight trials were combined, the pooled RR was 1.95 (CI 95%, 1.294–2.934; $p < 0.001$) showing that diagnostic stability of AD was higher than PDD-NOS (See Fig. 2). When looking at studies individually, Fig. 2 shows that all trials displayed a RR higher than 1.5 (demonstrating a higher diagnosis stability for AD than PDD-NOS) except

Chawarska et al. (2007) (RR = 0.987) who found a slightly lower diagnosis stability for AD than PDD-NOS.
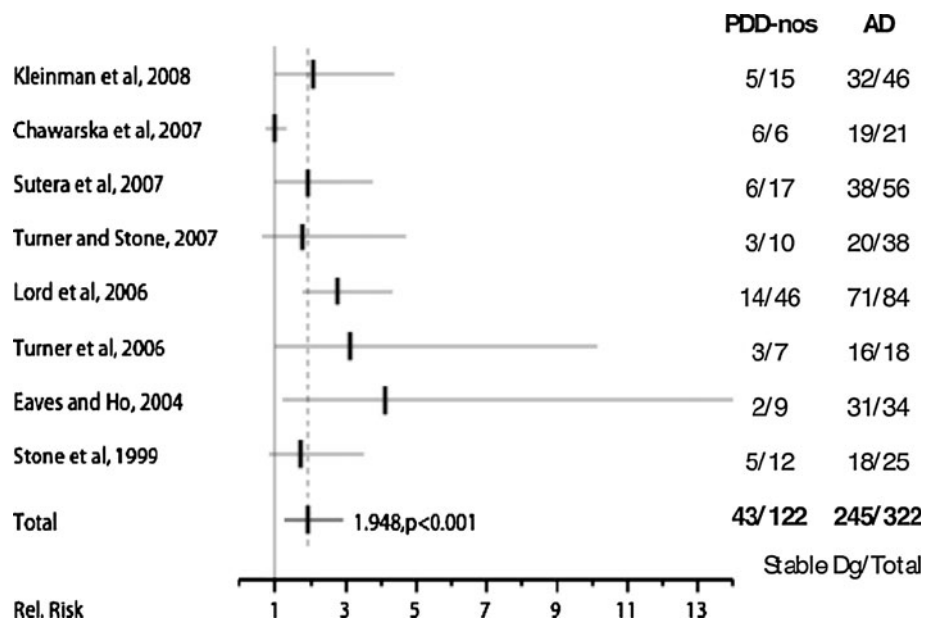
As all studies selected for the meta-analysis examined AD and PDD-NOS diagnoses before 36 months, this meta-analysis provides information about diagnostic stability in very young children. Reported diagnostic changes among the selected studies are summarized in Table 2. All studies except Chawarska et al. (2007) found that the stability of AD diagnoses were superior to 50% whereas PDD-NOS diagnosis was unstable over time, presenting with less than 45% stability in every study. In sum, 76% of the children diagnosed with AD before 36 months retained the same diagnosis whereas only 35% of PDD-NOS children did so over a mean period of 3 years.

Of note, two of the selected studies showed stability values which differed from the other studies. Over a 15-month inter-assessment period Chawarska et al. (2007) found a very high PDD-NOS stability over time (100%) in a very small sample ($n = 6$). Turner and Stone (2007) reported the weakest AD stability over time (53%). However it remained superior to the PDD-NOS stability over time reported in the same study (30%).

Developmental Trajectories for PDD-NOS

As diagnostic stability does not provide information on the evolution of children first diagnosed with PDD-NOS who did not retain the diagnosis, we examined the diagnostic trajectories between the two assessment times. Movements over time within the spectrum are presented in Fig. 3. After combining the 8 trials, we obtained 322 AD and 122 PDD-NOS children at Time 1. Of the 322 AD children at Time 1,

Fig. 2 Relative risk (RR) for diagnosis stability at T2 according to diagnosis at T1 (PDD-NOS vs. AD) (random model). *PDD-NOS* pervasive developmental disorder not otherwise specified, *AD* autistic disorder, *Stable Dg* number of participants with the same diagnosis at T1 and T2
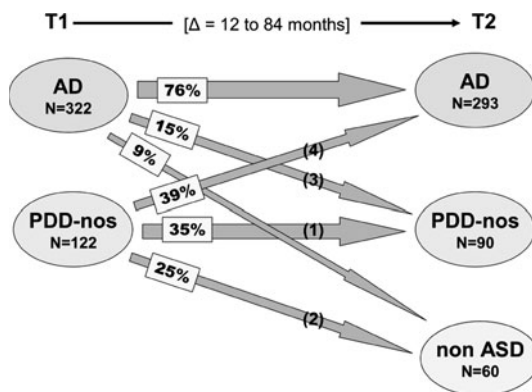
Fig. 3 Developmental trajectories within the autism spectrum. *AD* autistic disorder, *PDD-NOS* pervasive developmental disorder not otherwise specified, *non ASD* non autism spectrum disorder. Numbers within brackets refer to the discussion part of the text

245 remained AD (76%), 47 moved to PDD-NOS (15%) and 30 moved off the autism spectrum (9%). Of the 122 PDD-NOS at Time 1, 43 remained PDD-NOS (35%), 48 moved to AD (39%) and 30 moved off the autism spectrum (25%). In keeping with these findings the PDD-NOS condition seems to be almost equally dispatched over time to three statuses: (1) persistence of the same condition, (2) worsening of the phenomenology and (3) remission.

Discussion

We will first discuss findings related to the diagnostic stability, then the developmental trajectories of PDD-NOS over time, and finally the reliability of early PDD-NOS diagnoses in children less than 36 months.

PDD-NOS Diagnostic Stability

As it was mentioned in the Results section, the meta-analysis confirms the hypothesis that there is a higher diagnostic stability for AD than PDD-NOS (pooled RR of 1.95, $p < 0.001$). Pooling data from the selected studies indicated that an AD diagnosis was stable over time (76% stability) whereas PDD-NOS tended to be unstable over time (35% stability) (See Fig. 3).

A relatively low difference between AD and PDD-NOS stability was found in Turner and Stone (2007) who reported the lowest AD stability (53%). Table 2 shows that there are more AD children moving off the autism spectrum (31%) in Turner and Stone (2007), as compared to other studies. All children received speech therapy and the majority also received additional interventions (see Table 1). Thus, children from the Turner and Stone (2007) study could have shown a better improvement over time due to speech therapy which could in turn explain the lower

AD stability over time. This would be supported by Turner et al. (2006) who associated speech therapy with better outcome. It is also consistent with the evidence of possible benefits of early targeted intervention (Charman and Baird 2002). However such an assumption should be received with caution as many children in other studies received interventions during the inter-assessment period.

Unlike other studies where PDD-NOS stability was below 45%, Chawarska et al. (2007) elicited a very high stability (100%). This result could be explained by the very young sample of this study (Time 2 corresponded to age 3 years). Some studies reported that stereotyped behaviours, resistance to change and restricted interests seemed to appear later in children's development. These symptoms would be identified less consistently in 2 and 3 year old children (Charman and Baird 2002; Sutera et al. 2007; Kleinman et al. 2008). Thus, at 3 years of age, all symptoms are not necessarily displayed. This could explain why PDD-NOS diagnoses remained stable in this study. In addition, as for the previously discussed study, method considerations have to be taken into account. In Chawarska et al. (2007) inclusion derived from a selection process which combined information based on both the ADOS-G Module 1 and the ADI-R scored with the algorithm developed for children under 2 years. This could have led to the formation of more homogeneous groups recruited with more stringent criteria, which in turn fostered stability. Consequently high stability rates were elicited for both AD (90%) and PDD-NOS (100%). Moreover, at Time 2 the authors did not use the ADI-R and relied on observations drawn from the ADOS-G and choose a more elaborate module (Module 2) for 4 of 6 PDD-NOS participants to fit their speech development progress. Finally, the small interval between the two assessments (15 months) should be highlighted and seen as a key contributor for the high stability rates.

In sum, our meta-analysis confirms the hypothesis that AD bears a higher diagnostic stability than PDD-NOS which tends to be unstable over time (35% stability). With respect to this overall finding, the relative discrepancies found with two studies (Chawarska et al. 2007; Turner and Stone 2007) could be mainly explained by methodological differences affecting the design of the studies. Further studies on the stability of PDD-NOS should use ADI-R and ADOS at each diagnostic assessment time.

Developmental Trajectories for PDD-NOS

All studies included in the meta-analysis examined diagnostic stability in very young children (Time 1 corresponding to about 2 years of age in all studies). Two studies reported short-term stability from age 2 to age 3 years (Stone et al. 1999; Chawarska et al. 2007). These 2 studies reported

Table 1 Descriptive characteristics of the studies included in the current meta-analysis of diagnostic stability of AD and PDD-NOS overtime

Study	N	Sample	Age at T1: mean (SD)	Age at T2: mean (SD)	ΔT1-T2 (months)	Diagnostic procedure	Interventions during the study
Kleinman et al. (2008)	61	46 AD 15 PDD-NOS	27.1 months (6.4) 24.7 months (2.89)	51.1 months (6.8) 53.3 months (5.53)	24	Clinical judgment based on DSM-IV criteria ADI-R ADOS	Presence of interventions during the study
Chawarska et al. (2007)	27	21 AD 6 PDD-NOS	21.6 months (3.2) 21.6 months (2.5)	34.8 months (3.9) 38.1 months (8.3)	15	Best-estimate clinical diagnosis based on DSM-IV criteria modified for children under the age of 3, ADOS-G, ADI-R at time 1 only	Presence of treatment during the study
Sutera et al. (2007)	73	56 AD 17 PDD-NOS	27.6 months (4.7) 26.5 months (4.9)	52.2 months (6.6) 54.4 months (10.1)	24	Clinical judgment based on DSM-IV criteria ADI-R ADOS	All children received interventions
Turner and Stone (2007)*	48	38 AD 10 PDD-NOS	28.8 months (3.4) 28.8 months (3.4)	53.3 months (3.5) 53.3 months (3.5)	24	Clinical diagnosis based on DSM-IV or DSM-IV-TR criteria ADOS-G ADI-R at time 2 only	All children received speech therapy (mean hours per week:1.8) Majority of the children also received additional interventions
Lord et al. (2006)	130	84 AD 46 PDD-NOS	29.1 months (4.7) 29.1 months (5.6)	110.1 months (15.7) 114 months (17.1)	84	Best-estimate diagnosis including: Clinical diagnosis based on DSM-IV criteria, ADI-R, ADOS and the cognitive, language and adaptive test scores	Presence of treatment (with considerable variation in type and amount of treatment)
Turner et al. (2006)*	25	18 AD 7 PDD-NOS	31.0 months (3.8) 31.0 months (3.8)	108.8 months (7.9) 108.8 months (7.9)	84	Clinical diagnosis based on DSM-IV or DSM-IV-TR criteria ADOS-G ADI-R	Presence of intervention (speech-therapy and educational therapy)
Eaves and Ho (2004)	43	34 AD 9 PDD-NOS	33.0 months (4.58) 33.0 months (4.58)	59.0 months (7.47) 59.0 months (7.47)	24	Clinical judgment based on DSM-IV criteria	Presence of treatment (with wide variety in type and amount of intervention)
Stone et al. (1999)	37	25 AD 12 PDD-NOS	31.4 months (3.4) 31.4 months (3.4)	45.0 months (4.3) 45.0 months (4.3)	12	Clinical judgment based on DSM-IV criteria PL-ADOS at time 2 only	Presence of treatment during the study

* There was no overlap between the samples

AD autistic disorder, *ADI-R* Autism Diagnostic Interview-Revised, *ADOS* Autism Diagnostic Observation Schedule, *ADOS-G* Autism Diagnostic Observation Schedule-Generic, *CARS* childhood autism rating scale, *DAS* differential abilities scales, *DSM-IV* diagnostic and statistical manual of mental disorders-4th edition, *MSEL* mullen scale of early learning, *PDD-nos* pervasive developmental disorder not otherwise specified, *PL-ADOS* prelinguistic autism diagnostic observation schedule, *VABS* vineland adaptive behavior scales, *WISC-III* wechsler intelligence scale for children-3rd edition, *WPPSI-R* wechsler preschool and primary scale of intelligence-revised

Table 2 Diagnostic changes between Times 1 and 2 evaluations

Study	N	% AD Stability T1-T2	% PDD-NOS Stability T1-T2	Dx T1	Dx T2	% Dx Changes T1-T2
Kleinman et al. (2008)	61	70%	33%	AD	PDD-NOS	15
				AD	NON ASD	15
				PDD-NOS	AD	13
				PDD-NOS	NON ASD	53
Chawarska et al. (2007)	27	90%	100%	AD	PDD-NOS	9
				AD	NON ASD	0
				PDD-NOS	AD	0
				PDD-NOS	NON ASD	0
Sutera et al. (2007)	73	68%	35%	AD	PDD-NOS	21
				AD	NON ASD	10
				PDD-NOS	AD	24
				PDD-NOS	NON ASD	41
Turner and Stone (2007)*	48	53%	30%	AD	PDD-NOS	15
				AD	NON ASD	31
				PDD-NOS	AD	10
				PDD-NOS	NON ASD	60
Lord et al. (2006)	130	85%	30%	AD	PDD-NOS	14
				AD	NON ASD	1
				PDD-NOS	AD	58
				PDD-NOS	NON ASD	10
Turner et al. (2006)*	25	89%	29%	AD	PDD-NOS	0
				AD	NON ASD	11
				PDD-NOS	AD	43
				PDD-NOS	NON ASD	14
Eaves and Ho (2004)	43	91%	22%	AD	PDD-NOS	6
				AD	NON ASD	3
				PDD-NOS	AD	56
				PDD-NOS	NON ASD	22
Stone et al. (1999)	37	72%	42%	AD	PDD-NOS	24
				AD	NON ASD	4
				PDD-NOS	AD	50
				PDD-NOS	NON ASD	8
Total of 8 studies	444	76%	35%	AD	PDD-NOS	15
				AD	NON ASD	9
				PDD-NOS	AD	39
				PDD-NOS	NON ASD	25

AD autistic disorder, *PDD-NOS* pervasive developmental disorder not otherwise specified

the highest PDD-NOS stabilities (respectively 42 and 100%) as compared to other studies. As we previously mentioned, stereotyped behaviors, resistance to change and restricted interests seemed to appear later in children's development (Charman and Baird 2002; Sutera et al. 2007; Kleinman et al. 2008). Thus, at 3 years of age, all symptoms are not necessarily displayed. This could explain why PDD-NOS diagnosis is high. They also reported high AD stabilities (respectively 72 and 90%).

Four studies reported stability from age 2 years to age 4 years (Eaves and Ho 2004; Turner and Stone 2007;

Sutera et al. 2007; Kleinman et al. 2008). These four studies reported low PDD-NOS stabilities (respectively 22, 30, 35 and 33%). Stereotyped behaviors, resistance to change and restricted interests seemed to emerge at 4 or 5 years of age, which can lead to the change from PDD-NOS to AD diagnosis and could explain the low stabilities (Kleinman et al. 2008; Sutera et al. 2007). They also reported relatively high AD stabilities (respectively 91, 53, 68 and 70%).

Finally, two studies reported long-term stability from age 2 to age 9 years (Lord et al. 2006; Turner et al. 2006). These 2

studies reported high AD stabilities (respectively 85 and 89%) and low PDD-NOS stabilities (respectively 30 and 29%) over time. The above findings demonstrated that AD tends to be a stable and reliable diagnosis in 2 years old children and that PDD-NOS tends to be an unstable diagnosis in young children when established before 36 months.

From a developmental perspective, Fig. 3 displays the trajectories of the PDD-NOS conditions within the spectrum over time (12–84 months). In terms of development, PDD-NOS appears to be distributed along fourth situations (see Fig. 3): (1) A clinical construct per se likely depicting a subgroup of children with autistic features (35% retained the same diagnosis at Time 2), (2) A developmental disorder, e.g. communication delay, that is not so pervasive in some patients (25% are moving off the autism spectrum at Time 2), (3) A possible evolution of AD in some cases that show improvement (15% moved from AD to PDD-NOS), (4) A provisional diagnosis awaiting full AD criteria to be met (39% are moving from PDD-NOS to AD). The relevance of the fourth pathway could be supported by some studies which have reported that stereotyped behaviors, resistance to change and restricted interests seemed to appear later in children's development. The emergence of these behaviors by Time 2 can lead to the change from PDD-NOS to AD diagnosis (Kleinman et al. 2008; Sutera et al. 2007). In keeping with those assumptions illustrated in Fig. 3, PDD-NOS would be conceived as corresponding to a group of heterogeneous pathological conditions including prodromic forms of later AD, remitted or less severe forms of AD, and developmental delays in interaction and communication. Further studies on the stability of PDD-NOS should provide adequate information on the participants' gender, IQ, verbal and motor skills.

Reliability of PDD-NOS Diagnoses Before 36 Months

Towbin (2005) raised concerns about the validity of the PDD-NOS category because the criteria for this diagnostic category were vague and the group was heterogeneous. The DSM-IV-TR states that the PDD-NOS category "should be used when there is a severe and pervasive impairment in the development of reciprocal social interaction associated with impairment in either verbal and nonverbal communication skills, or with the presence of stereotyped behavior, interests, and activities, but the criteria are not met for a specific Pervasive Developmental Disorder" (APA 2000). Even if requiring impairments in at least two among the three domains, PDD-NOS remains an exclusion category used when the patient's phenomenology does not meet criteria for Asperger's Disorder or Autistic Disorder. In the same vein, Witwer and Lecavalier (2008) who cautiously examined the discriminant validity of ASD subtypes, found that the only difference between PDD-NOS and AD was the symptom

frequency, as PDD-NOS patients presented fewer symptoms in all three core domains when compared with AD. This finding is basically a tautology because it only follows the definition of the PDD-NOS category. The authors failed to find any external validators for the PDD-NOS category other than anxiety symptoms. Moreover they requested to interpret such a finding with caution because this difference might be due to differences in IQ between the studied groups. In addition the inter-rater agreement for PDD-NOS was found to be the lowest among ASD diagnoses (Mahoney et al. 1998). In sum, there is little support in the literature for the discriminant and inter-rater reliabilities of the PDD-NOS category.

Our meta-analysis confirms the hypothesis that AD bears a higher diagnosis stability than PDD-NOS which tends to be unstable over time with a 3-year stability rate of 35%. Despite two studies reporting high short-term stability rates for PDD-NOS between 2 and 3 years of age, long-term stability studies (from age 2 to age 9) demonstrated that PDD-NOS tends to be an unstable diagnosis in young children when established before 36 months. Therefore the literature seems not to support the discriminant and predictive validity of the PDD-NOS category in young children. Such a finding has clinical implications. Children whose PDD-NOS diagnosis was established before 36 months should be re-assessed at a later age. Given that PDD-NOS when diagnosed in very young children, covers a wide variety of pathological conditions including prodromic forms of later AD, remitted or less severe forms of AD, and developmental delays in interaction and communication, PDD-NOS should be considered at most as a provisory diagnosis, requiring reassessment at a later age.

With respect to the current proposed revision of DSM-IV-R criteria for ASD including merging the three subtypes, AD, AS and PDD-NOS, into one category, namely Autism Spectrum Disorder (ASD), our study confirmed the lack of support for reliably distinguishing PDD-NOS as a diagnostic entity. However our study clearly demonstrated the heterogeneity of the PDD-NOS group which will impact the predictive validity of the proposed DSM-V entity. The international clinical and research consensus on the robustness of AD as defined by currently more stringent DSM criteria, would be lost. International communication and comparison between interventions will be jeopardized.

Limitations

Several limitations to this meta-analysis should be noted. First, for the majority of the studies, clinical diagnosis at follow-up was not independent from initial diagnosis. Kleinman et al. (2008) suggested that a truly blind evaluation at Time 2 would be preferable because it would reduce potential bias in the diagnostic determination. However, Stone et al. (1999) evaluated the influence of

same and different clinicians on the stability of the diagnosis. They found no significant difference whether it was the same or a different clinician at Time 2. Further studies are needed to support this result. Second, general information that affects the evolution of AD and PDD-NOS were not documented. For example, the impact of the interventions on the final outcome was not measured, except in Turner et al. (2006). In addition there was not enough information to conduct subgroup meta-analysis based on gender ratios and IQ, as recommended for the latter by Witwer and Lecavalier (2008).

Conclusion

Our meta-analysis conducted on the eight longitudinal studies on PDD-NOS that have been published from 1996 to 2009 showed that PDD-NOS diagnosis was less stable than AD diagnosis. When established before 36 months, the overall stability rate was 35% at 3-year follow-up. Consistent with the previous literature on the reliability of the PDD-NOS diagnosis in young children, our meta-analysis did not to support the discriminant and predictive validity of this category. Thus, from a clinical standpoint, children whose PDD-NOS diagnosis was established before 36 months should be re-assessed at a later age. In addition further studies on the stability of PDD-NOS should use ADI-R and ADOS at each diagnostic assessment time and provide adequate information on the participants' gender, IQ, verbal and motor skills.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revisions ed.). Washington, DC: American Psychiatric Association.

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry, 43*(3), 289–305.

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J.-A., & Baird, G. (2005). Outcome at 7 years of children diagnosed with autism at age 2: Predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry, 46*(5), 500–513.

Chawarska, K., Klin, A., Paul, R., & Volkmar, F. (2007). Autism spectrum disorder in the second year: Stability and change in syndrome expression. *Journal of Child Psychology and Psychiatry, 48*(2), 128–138.

Cox, A., Charman, T., Baron-Cohen, S., Drew, A., Klein, K., Baird, G., et al. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry, 40*(5), 719–732.

Cucherat, M., Boissel, J. P., & Leizorovicz, A. (1997). EasyMA: A program for the meta-analysis of clinical trials. *Computer Methods and Programs in Biomedecine, 53*, 187–190.

Eaves, L. C., & Ho, H. H. (2004). The very early identification of autism: Outcome to age 4½–5. *Journal of Autism and Developmental Disorders, 34*(4), 367–378.

Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry, 66*(suppl. 10), 3–8.

Kleinman, J. M., Ventola, P. E., Pandey, J., Verbalis, A. D., Barton, M., Hodgson, S., et al. (2008). Diagnostic stability in very young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 606–615.

Klin, A., Lang, J., Cicchetti, D. V., & Volkmar, F. R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders, 30*(2), 163–167.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36*(8), 1365–1382.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*, 694–701.

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule-WPS edition*. Los Angeles, California: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

Mahoney, W. J., Szatmari, P., MacLean, J. E., Bryson, S. E., Jones, M. B., & Zwaigenbaum, L. (1998). Reliability and accuracy of differentiating pervasive developmental disorder subtypes. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*, 278–285.

Matson, J. L., & Boisjoli, J. A. (2007). Differential diagnosis of PDDNOS in children. *Research in Autism Spectrum Disorders, 1*, 75–84.

Moore, V., & Goodson, S. (2003). How well does early diagnosis of autism stand the test of time? *Autism, 7*(1), 47–63.

Saint-Georges, C., Cassel, Rs., Cohen, D., Chetouani, M., Laznik, Mc., Maestro, S., et al. (2010). What studies of family home movies can teach us about autistic infants. *Research in Autism Spectrum Disorders, 4*, 355–366.

Scahill, L. (2005). Diagnosis and evaluation of pervasive developmental disorders. *Journal of Clinical Psychiatry, 66*(Suppl 10), 19–25.

Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., et al. (1999). Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychology and Psychiatry, 40*(2), 219–226.

Sutera, S., Pandey, J., Esser, E. L., Rosenthal, M. A., Wilson, L. B., Barton, M., et al. (2007). Predictors of optimal outcome in toddlers diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 98–107.

Towbin, K. E. (2005). Pervasive developmental disorder not otherwise specified. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders Vol. 1: Diagnosis, development, neurobiology, and behaviour* (3rd ed., pp. 165–200). Hoboken: Wiley.

Turner, L. M., & Stone, W. L. (2007). Variability in outcome for children with an ASD diagnosis at age 2. *Journal of Child Psychology and Psychiatry, 48*(8), 793–802.

Turner, L. M., Stone, W. L., Pozdol, S. L., & Coonrod, E. E. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism, 10*(3), 243–265.

Van Daalen, E., Kemner, C., Dietz, C., Swinkels, S. H. N., Buitelaar, J. K., & Van Engeland, H. (2009). Inter-rater reliability and stability of diagnoses of autism spectrum disorder in children identified through screening at a very young age. *European Child and Adolescent Psychiatry, 18*, 663–674.

Walker, D. R., Thompson, A., Zwaigenbaum, L., Goldberg, J., Bryson, S. E., Mahoney, W. J., et al. (2004). Specifying PDD-NOS: A comparison of PDD-NOS, Asperger Syndrome, and Autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(2), 172–180.

Witwer, A. N., & Lecavalier, L. (2008). Examining the validity of autism spectrum disorder subtypes. *Journal of Autism and Developmental Disorders, 38*, 1611–1624.