

The ESSENCE-Questionnaire for Neurodevelopmental Problems – A Swedish School-Based Validation Study in 11-Year-Old Children

Valdemar Landgren^{1,2}, Leif Svensson³, Rajna Knez^{1,3}, Michail Theodosiou¹, Christopher Gillberg¹, Elisabeth Fernell¹, Magnus Landgren¹, Mats Johnson¹

¹Gillberg Neuropsychiatry Centre, Sahlgrenska Academy, Gothenburg University, Gothenburg, Sweden; ²Department of Psychiatry, Skaraborg Hospital, Skövde, Region Västra Götaland, Sweden; ³Department of Pediatrics, Skaraborg Hospital, Mariestad, Region Västra Götaland, Sweden

Correspondence: Valdemar Landgren, Department of Psychiatry, Skaraborg Hospital, Löfvängsvägen, Skövde, Region Västra Götaland, 541 42, Sweden, Email valdemar.landgren@gu.se

Purpose: To determine the prevalence of parent-rated developmental concern using the ESSENCE-Q (Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations-Questionnaire, 12-items, score range 0–24) and to ascertain the predictive validity and optimal cutoff level of the instrument in a school-based sample of 11-year-old children.

Methods: In a cross-sectional, school-based study, participants underwent a clinical assessment by a physician and a psychologist, teachers and parents completed the SDQ (Strength and Difficulties Questionnaire), medical health records and national tests were reviewed, and parents independently completed the ESSENCE-Q. In a case-conference outcomes were defined as a) the need for further clinical work-up due to suspected neurodevelopmental problems (NDPs) and b) degree of investigator-rated symptoms/impairment from NDPs on the CGI-S (Clinical Global Impression-Severity instrument, range 1–7, 4–7 defined as clinically symptomatic). Classification and optimal cutoffs of the ESSENCE-Q were determined using ROC (Receiver Operating Characteristic) analysis.

Results: Out of 343 eligible children, 223 enrolled, of whom 173 (50% of all eligible) had a parent-rated ESSENCE-Q. At least one of the 12 possible concerns was reported by parents of 36% of participants. Overall, in 101 (57%) participants a work-up was warranted, and 64 (37%) were clinically symptomatic from NDPs. The AUC of the ESSENCE-Q in detecting need for work-up was 0.70 (95% confidence interval [CI] 0.63–0.77), and the AUC in detecting clinically symptomatic participants was 0.82 (95% CI 0.76–0.88). ESSENCE-Q ratings correlated positively with CGI-S scores ($r=0.48$, $p<0.05$). A cutoff of ≥ 3 had the highest accuracy (78%) with a negative predictive value of 82%. Ratings >6 conferred few false positives cases with positive likelihood ratios >10 and positive predictive values of 86% or more.

Significance: This study of the ESSENCE-Q in 11-year-old children suggests it might be an acceptable instrument for screening of NDPs in children in middle school, optimally in conjunction with other methods.

Keywords: screening, neurodevelopment, school, ESSENCE-Q

Introduction

Neurodevelopmental disorders (NDDs) and neurodevelopmental problems (NDPs) are largely determined by genetics and prenatal factors, and their presentations and severity heterogeneous.¹ Many screening-tools have been developed for specific NDDs, but very few have been aimed at covering the full spectrum of NDDs/NDPs.^{2,3}

Early NDD symptoms are often not specific for one category of disorders (eg toddler temper tantrums may be related to impulsivity in attention-deficit/hyperactivity disorder [ADHD], rigidity in autism spectrum disorder [ASD] or simply

to psychosocial stressors). Symptoms from different diagnostic categories overlap, sometimes to the extent that unequivocal categorization may not be possible. For these reasons there is clearly a rationale for using broad screening tools.

Symptom levels and degree of impairment from NDDs vary across the life-span both in relation to maturational change and the level of environmental demands.⁴ Although early symptoms of NDDs can usually be ascertained in retrospect, the clinical presentation “as a case” is not limited to the preschool years, but extends into school-age period and even adulthood.⁵ Impairing symptoms most commonly emerge in early school years, the age at which the incidence of NDD diagnoses peaks in for example Sweden and Denmark.^{6–8} Since children often present for evaluation in school-age years, broad screening for NDDs during middle school may provide a means to identify a second wave of “onset” cases (those whose symptoms reach a clinical level only at this stage). The term ESSENCE (Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations) was coined by Gillberg, to emphasize the early onset and non-specificity of signs of developmental deviations and problems that may later be categorized under specific diagnoses, where co-occurrence is the rule rather than exception (Included under the ESSENCE umbrella are: ADHD, ASD, oppositional defiant disorder [ODD], developmental coordination disorder [DCD], Tic disorder/Tourette syndrome [TS], language disorder [LD], dyslexia, dyscalculia, intellectual disability [ID], borderline intellectual functioning [BIF], behavioral phenotype syndromes [eg Rett syndrome or 22q11-deletion syndrome, fetal alcohol syndrome and other “etiological” syndromes] and also epilepsy and cerebral palsy, that may present to the clinic with NDD symptoms) (Gillberg 2010). The introduction of the ESSENCE-concept was followed by the development of the ESSENCE-questionnaire (ESSENCE-Q) intended for use as screening tool covering the whole spectrum of NDDs. It is currently used in clinical and research contexts and has been validated in population-based and clinical samples in Japan and Europe, mostly in pre-school aged children.^{9–11} One study investigated teacher-rated ESSENCE-Q in Swedish school-aged children, but validation was restricted to intellectual functioning, showing that BIF was associated with higher teacher ESSENCE-Q-scores.¹² No school-based studies have established its validity in relation to other clinical-level problem groups.

Aim

Our aims were to (i) determine the prevalence of parent-rated developmental concern using the ESSENCE-Q, and (ii) ascertain the predictive validity for impairing NDD symptoms and optimal cutoff level of the instrument in a non-clinical sample of 11-year-old children.

Methods

Setting and Participants

This was a cross-sectional study conducted in regular schools in western Sweden between 2018 and 2020 as part of a broad research initiative covering school children’s general health and abilities, and their relations to school performance. A convenience sample of six schools was selected based on the authors’ knowledge of local headmasters’ readiness to engage in studies of children’s health. Parents of all pupils attending 4th grade at these schools were invited to participate. At each school site, 1–2 academic years of children were targeted, and participation constituted an add-on to the existing routine health check-up in 4th grade. The targeted schools approached had no pupils attending special education classrooms. A flow diagram of participants and measures is shown in [Figure 1](#). To reduce bias, all assessors and informants (parent, teacher, maternal interviewer, physician, psychologist) provided data independently of each other with no other clinical information available to them at the time of assessment. After participating parents had received questionnaires together with the consent form and returned them in pre-stamped envelopes, physical neurodevelopmental in-school examinations were scheduled, followed by neuropsychological assessment of the child, requisition of medical health records, and maternal interviews. In some instances where consent forms but not questionnaires had been returned, reminders were sent out to parents in connection with the maternal interviews. Lastly, a clinical synopsis of all available information was performed in a case-conference (described below). In a few instances where substantial clinical needs had been discovered (after face-to-face assessments or the case-conference), referrals to appropriate regular care services were made in liaison with the school nurse and parents.

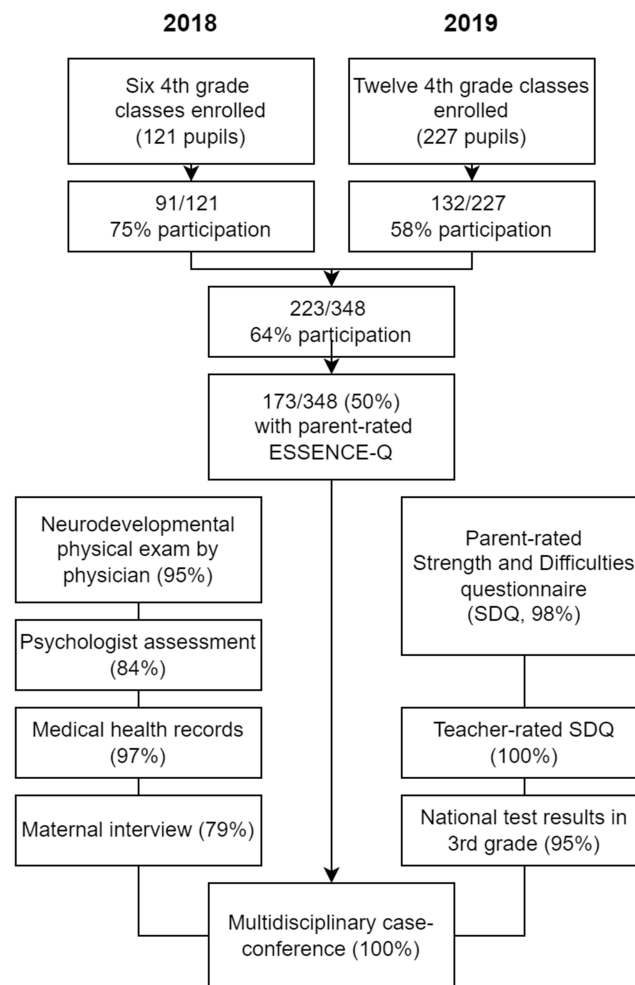


Figure 1 Flow diagram of participants and measures used in the school-based study of 11-year-old children.

Measures

Exposures

We investigated the capacity of the parent-rated ESSENCE-Q for detecting clinically relevant parent concern. The ESSENCE-Q is not intended for establishing a diagnosis, but to be used as a screening tool for the identification of children who might benefit from further assessment of functional areas negatively affected in NDDs, or in other words for identifying those who are not in need of further assessment. The ESSENCE-Q consists of 12 items referring to different developmental/behavioral domains (general development, motor development and milestones, sensory reactions, communication, language/babbling, activity/impulsivity, attention, social interaction/interest in other children, behavior, mood, sleep, feeding and absences/other types of paroxysmal attacks), individual items rated as no problem =0, maybe, or a little =1, Yes=2 (range 0–24). Parents filled out the questionnaire at home and returned them via post in a pre-stamped envelope, or could choose to fill out the questionnaire in a data-protected web-based form. To avoid bias, they had not received any information about other assessments in the study prior to completing questionnaires or instances with reminders when being interviewed.

Outcome

To comprehensively evaluate a child presenting with developmental concerns a full assessment is needed, including history-taking regarding pre- and perinatal risks, possible associated medical conditions, family history regarding NDDs, detailed structured and recorded clinical observation and examination on site, and a review of possible psychosocial

stressors/buffers - all against the background of environmental demands and the child's individual trajectory of maturation.

To take this complexity into account and improve clinical interpretability of the ESSENCE-Q, we defined two outcome levels, (i) clinician's decision of need for further clinical evaluation, and (ii) clinician's impression of overall severity/impairment. Outcome level assignment was made in case-conferences on the basis of results obtained from seven independent sources of information; medical health records, neuropsychological assessment, neurodevelopmental/physical examination, maternal interview, parent- and teacher rating scales, and national tests (described below).

Outcome Specification

For the *first outcome level*, a clinical synopsis of all available data for each participant was made at a case-conference. Consensus decisions were made regarding whether or not symptoms/impairments signaled a need for clinical work-up regarding NDP/NDD. The term NDP refers to deficits/difficulties in specific intellectual, motor, language or social functions with onset during the developmental period, causing impairments or distress irrespective of whether or not a diagnostic "cutoff" is met. In cases meeting a cutoff, the problems would render a diagnosis of NDD. The first outcome level was operationalized as follows: 0 = No indication of NDPs, 1 = Some indication that NDPs cannot be ruled out, but is probably maturational, unlikely to be of clinical significance and does not warrant further evaluations, 2 = Clinical picture may be consistent with NDPs and warrants further work-up, 3 = Full clinical work-up indicated, NDPs consistent with one or more specific NDDs. A rating of 2 or 3 in this system did not differentiate severity, only whether the clinical picture unequivocally was consistent with a NDD (scored as 3) or not.

At the *second outcome level*, with a view to evaluate whether or not ESSENCE-Q ratings were related to symptom level and/or degree of impairment, all participants were rated on the Clinical Global Impression – Severity instrument (CGI-S, range 1–7).¹³ This method was chosen because assessing all participants was considered important for establishing validity of a screening measure, although no informant interview formally assessing diagnostic criteria for each NDD could be performed. The CGI-S reflects the clinician's impression of degree of impairment with reference to the distribution of the specific condition under study, based on all available information about the participant. In the present study, we adapted the scale as follows: 1=no indications of NDPs; 2=some symptoms but negligible impairment unlikely to amount to clinical level; 3=some symptoms likely accounting for a level of impairment that would benefit from further work-up, but not necessarily amounting to a clinical NDD diagnosis; 4=obvious symptoms, either subsyndromal across several functional areas or likely amounting to one clinical NDD diagnosis with mild impairment; 5=obvious symptoms, likely amounting to at least one clinical diagnosis with moderate impairment; 6=severe symptoms amounting to more than one clinical NDD diagnosis with severe impairment reflected in difficulties with academic achievement and daily functioning requiring extensive support from other than parents only; 7=very severely impaired with multiple NDDs. In a clinical setting based on our experience, participants with CGI 1–3 would not be referred for work-up in regular health care (child psychiatry), whereas participants with CGI 4–7 would likely have symptoms/impairments warranting clinical attention and likely diagnosis. To exemplify: i) some coordination deficits in the physical exam and a positive Stroop test in neuropsychological testing (which may indicate problems with attention on mundane tasks), but no evident behavioral or academic problems would be rated as a "1" (slight NDPs, unlikely significant), and a CGI-S of 2 (few non-impairing symptoms), ii) history of febrile seizures, coordination problems in the physical exam, high ratings in hyperactivity/inattention domain from teachers and parents, monthly tension headaches requiring medication, but no academic problems would be rated as a "2" (Possible NDPs, work-up indicated), and a CGI-S of 3 (some impairment, not necessarily at a clinical level), and iii) attention deficits in both psychological and physical exam, internalizing symptoms, frequent headaches, fine- and gross motor coordination problems, high behavioral ratings in hyperactivity/inattention from teachers and parents would be rated as a "3" (work indicated and consistent with specific NDD) and a CGI-S of 5 (evident impairment, at least one clinical diagnosis).

All case-conferences were attended by a child neurologist (ML) and neuropsychologist (LS), both with more than 25 years of clinical experience working with families with children with NDDs, and a psychiatrist (RK or VL). Where information was considered sufficient, descriptive symptomatic areas were assigned comprising externalizing (symptoms of ADHD, ODD, CD), autism spectrum (social interaction deficits), speech-language or learning (indications of language

problems, dyscalculia, dyslexia) and developmental coordination (gross motor deficits). Reliability of the case-conference results was established against three other clinicians with extensive experience of NDDs (a clinical professor [CG], and an associate professor of child psychiatry [MJ] and a child neurologist [MT]) who independently rated 11 cases.

Data Sources

Medical Health Records

Almost all Swedish children attend child health care check-ups from birth up to 5 ½ years of age, comprising about 14 visits to a nurse and at least 2 visits to a physician. The check-ups are intended for developmental and anthropometric assessments and vaccinations.

Upon school entry, responsibility is transferred from child health care to the school health care, where a nurse continues monitoring anthropometric measurements, vaccinations and at least two general health check-ups with interviews during primary school, and as needed. In the present study, child health, school health, and associated regular medical health records were reviewed for signs of developmental deviations (eg, raised concerns, referrals, descriptions, diagnoses) and taken into consideration.

Psychological Assessment

Participants were assessed with the Leiter-III non-verbal intellectual quotient (IQ) test by an experienced psychologist, taking 60 minutes.¹⁴

Physician-Led Physical Assessment

The physician performed a physical exam with emphasis on motor performance, interaction and dysmorphology and brief questions about general health, for 20 minutes, all according to a standardized protocol, with a scribe taking notes.

Maternal Interview

Nurses were recruited and trained in performing interviews, according to a standardized form with emphasis on nutritional factors for about 45 minutes.

Teacher and Parent Ratings

Teachers and parents independently filled out the strength and difficulties questionnaire (SDQ) with 25 items comprising a total score (range 0–40) and five subscales (range 0–10); emotional symptoms, conduct problems, hyperactivity-inattention, peer problems, and prosocial behavior (prosocial behavior is not included in the total score).¹⁵

National Tests

Results from national tests in Swedish and mathematics were reviewed. On a national level, the success rate in the eight subtests in the Swedish language range from 90% to 98%, with boys faring worse than girls.¹⁶ The proportion of students passing the seven mathematical subtests range from 83% to 96%, with no major differences by gender. About 2–3% of students are considered unfit to take the tests because they are deemed to lack sufficient skills.¹⁷ Non-participation and number of failed tests were therefore both taken into consideration.

Statistics

Based on prior studies on the prevalence of NDDs in similar contexts,^{18–20} the expected prevalence of clinically significant NDPs was set to 10%. We considered an area under the curve (AUC) of 0.75 to be a minimally acceptable performance of the ESSENCE-Q in separating those with a clinical level of NDPs from those without. According to the power calculation, an AUC of 0.75 with alpha set to 0.05 and beta to 0.80, required a sample size of 150 (Xavier 2011). We employed no imputation procedure for missing data. In analyses, receiver operating characteristic (ROC) curves were established for the ESSENCE-Q summary score against two binary definitions; 1) cases with ESSENCE where a thorough clinical work-up would be indicated (ESSENCE 2 or 3 according to the first outcome level) vs those where it would not be indicated, and 2) Cases decided as having a level of impairment and/or symptoms reaching an undoubtedly clinical level (CGI level 4–7, second outcome level) vs those with scant signs, at most suggestive (CGI level

1–3). Alongside AUC-curves, we reported classification tables with each diagnostic cutoff, including sensitivity, specificity, accuracy (proportion accurately classified, [true positive + true negative/all positive and negative]) negative predictive value (NPV=true negative/all negative) positive predictive value (PPV= true positive/all positive), Youden index (sensitivity + specificity –1), positive (LR+) and negative likelihood ratio (LR-) ($LR+ = \text{True positive}/\text{False positive}$, $LR- = \text{True negative}/\text{False negative}$). Group comparisons were analyzed with non-parametric tests. All analyses were performed in R 3.6.3 (R Core Team 2020).

Ethics

All caregivers and children signed informed consent. The study was approved by the ethical review board at Gothenburg university in accordance with the Declaration of Helsinki (No. 852–17).

Results

School Characteristics and Non-Participation

Participants were recruited from eight schools (two private, six public) in five municipalities in Sweden. According to statistics from the Swedish National Agency for Education,²¹ participating schools had a socioeconomic index (based on parental education level, income and dependency on social benefits) in the lower range (mean percentile 29th, range 1st to 75th) compared with other schools in Sweden.²² There were no data for one school (n=11).

Attrition varied between the two recruited academic years (74% vs 54% participation), and between schools (range 36–91%). Out of 221 participants, 173 had parents providing an ESSENCE-Q and were included in the study, rendering total attrition for the current report at 50%. CGI-ratings did not differ between those with and without a parent-rated ESSENCE-Q (Wilcoxon test $p=0.36$). There was no data available for drop-out analysis of eligible children declining to participate. However, as reflected in the differing participation rates between schools and classes, attrition in study participation was deemed contingent on school headmasters' commitment to the project and their communication to parents. Pursuit of data completeness and ESSENCE-Q ratings was hampered by COVID-19 pandemic restrictions that made face-to-face follow-up of non-responders impossible. Data availability in this study (N=173) therefore ranged from 137 (79%) for the maternal interview, to 173 (100%) for SDQ-ratings (Figure 1).

Participant Characteristics and ESSENCE-Q Ratings

Three parent ratings were re-coded, because comments or other sources of information made it obvious that they had been misunderstood and coded in reverse (“Yes” was misunderstood as implying no problem). Thirteen participants had been rated twice, eight were specified as rated by both parents, two twice by the same mother and in three instances the relation of the rater was not reported. Participant characteristics and ESSENCE-Q summary score distributions are provided in Table 1. Non-verbal IQ of participants were in the normal range (range 77–132). As further depicted in Figure 2, ratings in the ESSENCE-Q were skewed, with a majority (56%) of parents scoring either “Yes” or “Maybe/a little” in one area or more, 48% scoring at least one area with “Maybe/a little”, and 36% scoring at least one area as “Yes”. Summary scores did not differ by gender. Concerns were most common in the developmental areas attention, activity/impulsivity, mood and feeding (Figure 3).

Outcome Specification

The reliability of the case adjudication was tested independently against the consensus of three experts who reviewed the same information (see section on outcome specification). Out of 11 cases, there was perfect agreement in six, one participant considered in need of work-up was downrated by the experts, and four were uprated. It was concluded that study assessors were slightly more restrictive in their definition as to whether NDPs were present than the independent experts. Of 173 participants in the study, 101 (57%) met the case-conference definition for a recommended work-up (clinical work-up likely reflecting NDPs to some extent is warranted) and 64 (37%) had a CGI-S of 4 or more (clinically relevant impairment, CGI-S 4 [n=40], 5 [n=20], 6 [n=2]). For 53 (30%) NDPs could be categorized as externalizing (20%), speech-language or learning (3%), autism spectrum (5%), or developmental coordination (8%). Prior to school

Table I Clinical Characteristics and Parent-Rated ESSENCE-Q

Participant Characteristics	Overall (N=173)	No Clinically Impairing NDPs (n=109)	Clinically Impairing NDPs (n=64)
Age at physical examination, mean (SD)	11.0 (0.4)	11.0 (0.3)	11.0 (0.4)
Boys, n (%)	99 (57)	57 (52)	42 (66)
Living with both parents	112 (70)	82 (80)	30 (53)
SDQ total difficulties score, teacher rated	6.9 (6.1)	4.2 (4.3)	11.6 (5.8)
SDQ total difficulties score, parent rated	8.7 (6.9)	5.9 (4.6)	13.8 (7.3)
ESSENCE-Q score ^a , m (SD)	3.6 (5.1)	1.4 (2.5)	7.2 (6.2)
Boys	3.7 (5.0)	1.5 (2.7)	6.5 (5.8)
Girls	3.5 (5.3)	1.3 (2.3)	8.5 (6.7)
ESSENCE-Q score distribution, n (%)			
0	75 (43)	66 (61)	9 (14)
1–2	28 (16)	20 (18)	8 (13)
3–6	37 (21)	18 (17)	19 (30)
7–24	33 (19)	5 (5)	28 (44)

Notes: Clinically impairing neurodevelopmental problems (NDPs) was defined as a Clinical Global Impression-Severity rating of 4–7 (range 1–7). Measures are reported as number (%) or mean (standard deviation). ^aFor participants rated by two parents (n=13) the average score was used.

entry, three participants (2%) had undergone a work-up for NDDs and one (1%) had regular care contact with habilitation service due to cerebral palsy. After school entry but before study enrollment, 16 participants (9%) had undergone a work-up for NDDs. Assuming there were no cases with clinically symptomatic NDPs among non-participants the results translate to a 19% (65/345) minimum prevalence of NDPs (but not necessarily corresponding to the level of a categorical NDD-diagnosis) in the whole cohort. ESSENCE-Q and CGI-S ratings were moderately correlated (Spearman's $r=0.48$, $p<0.01$, [Figure 4](#)), as were parent-rated SDQ total difficulties score and CGI-S (Spearman's $r=0.56$, $p<0.01$).

Classification Capacity of the ESSENCE-Q

In [Figure 5 A-B](#), ROC-curves of the classification capacity of the ESSENCE-Q as exposure, and the two outcome levels as outcome is depicted ([Figure 5A](#), clinical work-up for NDPs warranted, [Figure 5B](#) clinically symptomatic/impairing NDPs; CGI-S 4–7). The AUC of the ESSENCE-Q in predicting need for work-up was 0.70 (95% confidence interval [CI] 0.63–0.77), with the highest accuracy (Proportion correctly classified) obtained at a cutoff ≥ 1 (sensitivity 66%, specificity 63%, accuracy 65%, NPV 54% PPV 74%). In predicting clinically impairing NDPs (CGI-S rating ≥ 4) the AUC of the ESSENCE-Q was 0.82 (95% CI 0.76–0.88), with the optimal cutoff of ≥ 3 (sensitivity 81%, specificity 72%, accuracy 78%, NPV 82%, PPV 71%) (Classification [Table S1A](#) and [B](#) in [Supplemental Material](#)). By both definitions, an ESSENCE-Q summary score >6 conferred few false positive cases and had likelihood ratios ~ 10 or more ([Table S1A](#) and [B](#)). A cutoff of 0 had the highest NPV (87%) toward clinically impairing NDPs.

Discussion

In this school-based cohort of 11-year-old children, we found a high prevalence of parent developmental concern as rated with the ESSENCE-Q. There were indications for clinical work-up related to NDPs in half of the cohort, and clinically significant impairment in a third. The ESSENCE-Q summary score correlated with degree of impairment, with high scores conferring few “false positive cases” and a high likelihood ratio for clinically relevant problems, whereas low scores showed high NPVs.

The prevalence of clinically impairing NDPs is higher than commonly reported prevalence rates of NDDs, which may have several reasons. First, a majority of studies report NDD diagnoses from population registers, from screening-procedures using cutoffs that are statistically motivated, or strictly applying diagnostic criteria in an “either or”-fashion, and consequently arrive at smaller proportions clinically symptomatic, around 10–15%.^{2,18} When aggregating results studying specific

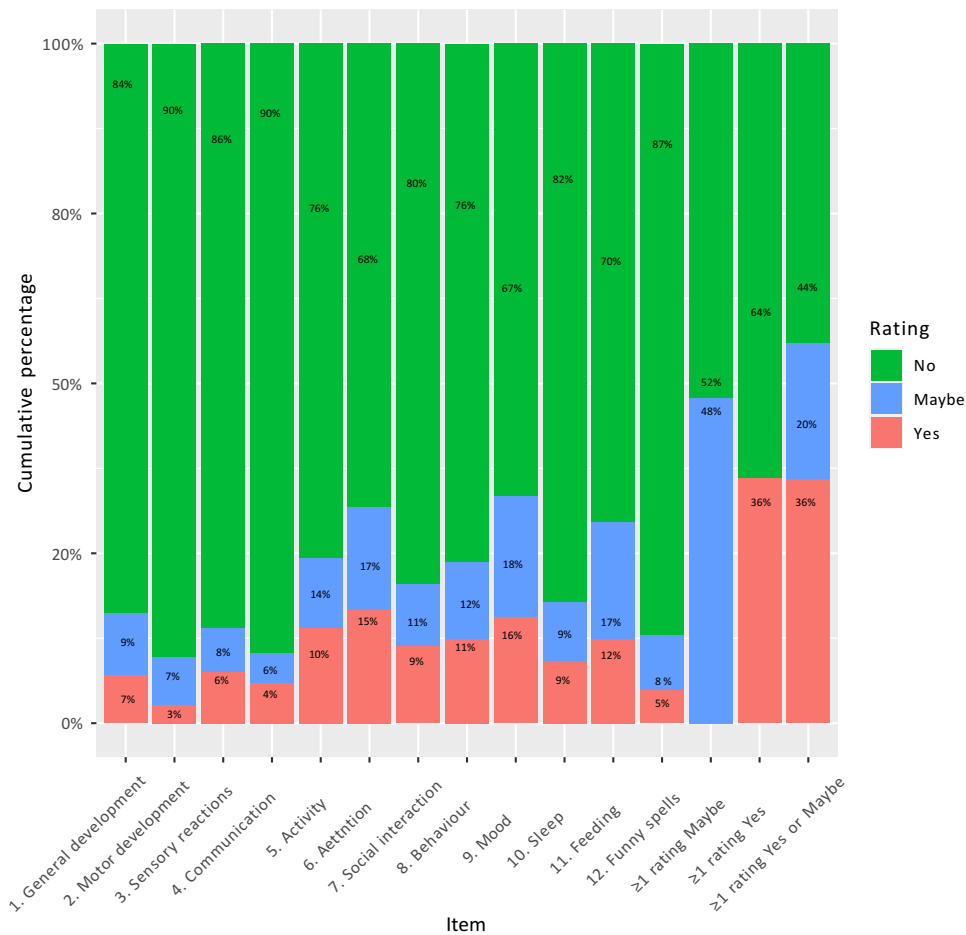


Figure 2 Item-wise distribution of parent-rated ESSENCE-Q in 11-year-old children (N=173).
Notes: Distribution of parents' answers to ESSENCE-Q items 1–12. Last three bars denote the occurrence of any of the twelve items rated with at least 1 "Maybe/a little", at least 1 "Yes" or at least 1 rating "maybe/a little" or "Yes".

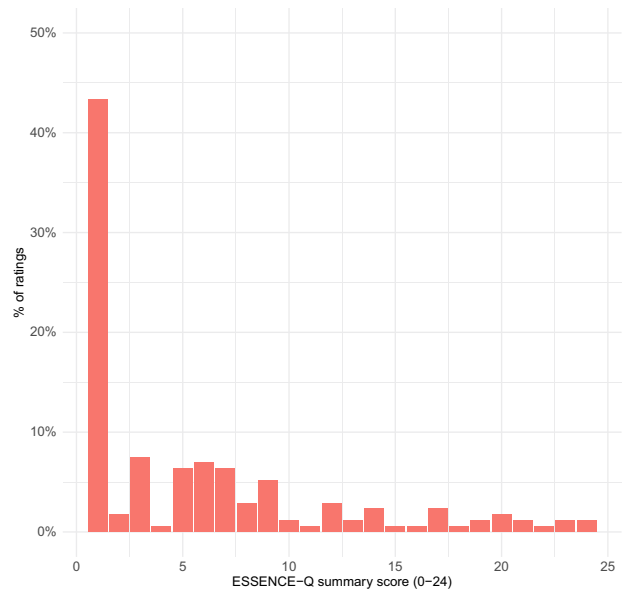


Figure 3 Distribution of parent-rated ESSENCE-Q summary scores in the school-based study of 11-year-old children (N=173).
Notes: The mean value was used for 13 participants rated by 2 parents.

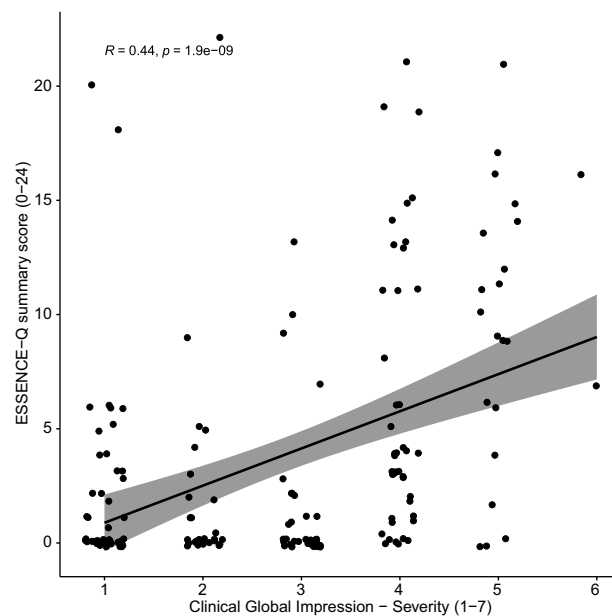


Figure 4 Correlation plot of ESSENCE-Q rating and Clinical Global Impression – Severity in the school-based study of 11-year-old children (N=173).
Notes: R denotes Spearman's rank correlation between the two variables.

disorders, subthreshold cases and those not ruled in for a diagnosis even though the aggregate of impairment across functional areas is impairing, may therefore have been overlooked. Second, there is evidence to suggest that although symptom prevalence across the population has been relatively stable during the last decades, less symptomatic individuals report more impairment than previously, which may increase the proportion with clinical-level problems.^{23,24} Third, early school years constitute the peak for effects from individual immaturity, as chronological age differences relative to peers can be substantial. It has been shown to increase the liability for impairing NDPs in this phase of life, although they are transient for many.²⁵ Fourth, the somewhat skewed sample as indicated by socioeconomic status (around 1 SD below the mean in Sweden) may indicate a somewhat increased familial liability for NDPs in this cohort.^{26,27} This is consistent with the higher SDQ-scores compared with those reported in a population-based sample of Swedish children.²⁸ As shown previously, SDQ-ratings from teachers tend to be lower than those obtained from parents.²⁹

Because a minority of participants with impairing NDPs had received a diagnosis of NDD prior to the study, it cannot be ruled out that parents of these children may have been more prone to participate. Even so, the proportion clinically impaired is not far from that reported in a previous Swedish school-based cohort with high participation rate, where 30% screened positive for behavioral and/or learning problems.^{30,31} Of all rated as clinically impaired the majority had a CGI-S = 4 (n=40), indicating the lowest level of clinical problems. The proportion with CGI-S ratings >4 (13% of participants, 7% of the eligible sample) are children who we considered similar to those diagnosed with NDDs in previous population-based studies.^{19,20} The relatively high rate of clinically impaired cases is in line with reports indicating that some children who in previous years would not have received a diagnosis, do so in current clinical practice.³² However overall, the most conservative prevalence (19%) is higher than expected. It is important to keep in mind here that NDP is not equivalent with NDD. It may be due to a combination of sample selection, and an environmental shift in the school environment that make “borderline cases” with NDPs clinically impaired.¹²

Classification accuracy (proportion accurately classified) as well as the Youden indices for the ESSENCE-Q imply that screening-procedures for NDPs based on parent concern only will be insufficient. “False negative” cases could be due to some parents perceiving problems as being due to medical conditions or psychosocial stressors and not to NDPs, or that study assessors may have identified problems that parents deemed irrelevant. “False positive” cases reflect that parent concern is more common in this age group than in preschool-aged children. Indeed, in a population-based study by

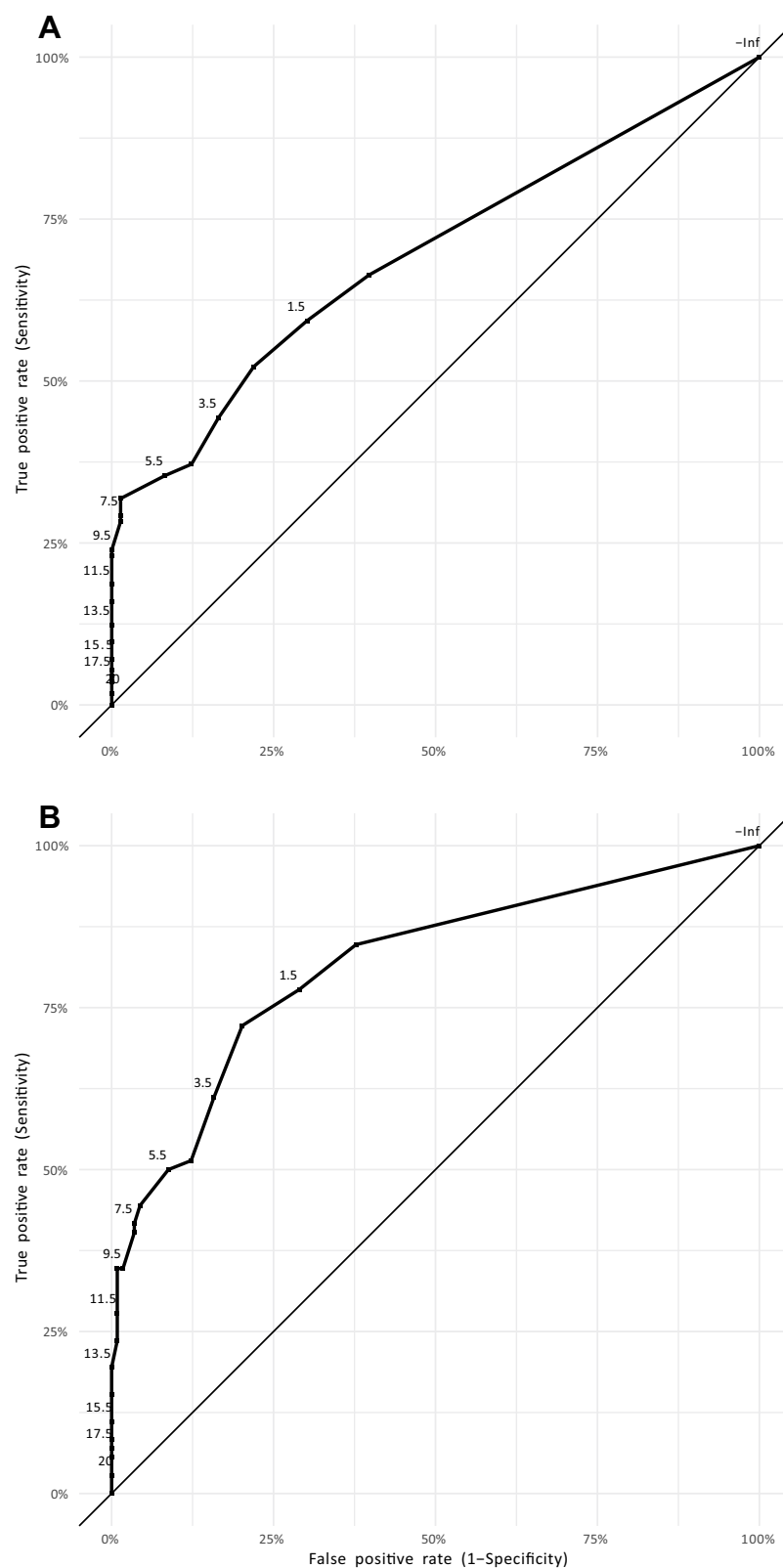


Figure 5 (A) Receiver operating characteristic curve of the classification accuracy of the ESSENCE-Q toward indication for further work-up related to neurodevelopmental problems to some extent in the school-based study of 11-year-old children. (B) Receiver operating characteristic curve of the classification accuracy of the ESSENCE-Q toward clinically impairing neurodevelopmental problems (a Clinical Global Impression-Severity of 4–7) in the school-based study of 11-year-old children.

Notes: Points and numbers denote selected cutoffs in the ESSENCE-Q. (A) Further work-up indicated in 101 of 173 participants. Area under the curve = 0.70 (95% confidence interval 0.63–0.77). (B) Clinically relevant neurodevelopmental problems found in 65 of 173 participants. Area under the curve = 0.82 (95% confidence interval 0.76–0.88).

Marinopoulou et al the mean ESSENCE-Q score from parents of 7-year-old children was 1.7, which is less than half of that found in this study.⁹

With regard to optimal cutoffs, there are studies of younger children for comparison. Hatakenaka et al investigated the validity of ESSENCE-Q in detecting NDDs in three studies of preschool-aged children, one cohort referred for a work-up for NDDs, and two cohorts attending regular public health check-ups.^{11,33,34} There was a high prevalence of NDDs in the studies of public health check-ups (9% and 17%). With the exception of ratings at 20 months of age, the AUC of ESSENCE-Q rated by mothers was low, whereas expert-rated ESSENCE-Q (by nurse or psychologist) showed good predictive validity with scores of 2 or 3 used as cutoff.

Kattimani et al investigated the predictive validity of caregiver-rated ESSENCE-Q for detecting NDDs in a hospital-based early intervention clinic (N=100, age 1–4 years) and found an AUC of 0.95, with 4 or more being the optimal cutoff.³⁵

Stevanovic et al translated and validated the ESSENCE-Q in a southern European multi-center study of children diagnosed with NDDs and a comparison group of children attending health-care for other reasons or ruled out as having NDDs as a comparison group.¹⁰ The AUC of the ESSENCE-Q in all forms (parent-completed form, telephone- or direct interview forms administered by trained nurse or specialist) was very high (0.91–0.96), with scores of 3–5 being optimal cutoffs (lower for parent-rated than expert-rated forms), although the sample selection indicate that the AUC may not be entirely transferable to a population-based setting.

We propose that the ESSENCE-Q in school-aged children would be best used for screening in conjunction with other methods, such as acting on information acquired in the school routine health check-ups and school health records. This is in line with a recent meta-analytic study of screening measures for childhood ADHD, concluding that one informant only is insufficient for optimal screening.² The likelihood ratios (LR) and predictive values on the high and low ends in the ESSENCE-Q (eg scores < 3 having LR- <0.5 and NPV >84% and scores >6 having LR+ >10 and PPV >86%) are in a range that may aid clinical decision-making. Therefore, if combined with other measures in a non-clinical screening setting, low ESSENCE-Q scores can assist in ruling out the need for further inquiry, and high ESSENCE-Q scores, exhibiting very high specificity, will likely identify “true” cases and provide an indication of the degree of impairment, which is supported by the correlation between parent concern and CGI-S ratings. This is in agreement with a study by Reilly et al involving preschool-aged children with epilepsy and a comparison group of clinic-attenders with NDDs, both groups markedly impaired, where the mean ESSENCE-Q score was around 12 for both groups.³⁶

Several study limitations should be noted. Generalizability is restricted to settings similar to those in the Swedish school health services, both as regards cultural notions about seeking help in health care services and as regards familiarity with the concept of NDDs. The sample size was modest, and attrition considerable. Whether responses for internet-administered ESSENCE-Q differ from the paper-and-pencil version has not been formally assessed here, but differences between paper-and-pencil and internet-administered versions of similar brief questionnaires has been negligible.²⁸ Outcome definitions were purposefully broad, but their validity over time is uncertain, especially in cases considered mildly affected (eg CGI-S of 4), where NDPs may be more transient and thus overstated. “Mild cases” may also on the contrary be understated and not encompass fully the scope of problems to be unfolded in subsequent years. The reliance of the CGI-S ratings may impact reproducibility of the findings herein, depending on the extent to which the clinical judgment of the study assessors is generalizable to other clinicians and researchers. We think prior research experience in the field, and operationalized CGI-S definitions mitigate this limitation to some extent, and that the use of the SDQ, reliability testing of the case-conference procedure provides some concurrent validity, but rater bias is not eliminated.

Conclusion

We found parent concerns as reflected by results obtained using the ESSENCE-Q in a school health context to be common, and to be correlated with severity of NDPs, and that high ESSENCE-Q scores in most cases identified clinically relevant problems. The study provides additional support for its use in screening. Future studies are needed

using larger samples of school-aged children and (hopefully) with higher participation rates. Focusing on the ESSENCE-Q, plus additional modes of screening used in conjunction, would be of particular interest.

Implications and Contributions

- The ESSENCE-Q as a screening measure for neurodevelopmental disorders has been validated previously in children of preschool age. However, many individuals with neurodevelopmental problems are detected at a later stage when attending school, and the validity of the ESSENCE-Q in this age-group is largely unknown.
- In a school-based cohort of 11-year-old-children, the AUC in detecting clinically symptomatic participants with neurodevelopmental problems was 0.82 (95% CI 0.76–0.88).
- Scores in the ESSENCE-Q correlated positively with clinical global impression severity of impairment ($r=0.48$, $p<0.05$).
- There was considerable attrition, with 50% of eligible children participating.
- The study provides additional support for the use of ESSENCE-Q in screening, optimally in conjunction with other methods.

Funding

Funding was provided by the research fund at Skaraborg Hospital.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Thapar A, Cooper M, Rutter M. Neurodevelopmental disorders. *Lancet Psychiatry*. 2017;4(4):339–346.
2. Mulraney M, Arrondo G, Musullulu H, et al. Systematic review and meta-analysis: screening tools for attention-deficit/hyperactivity disorder in children and adolescents. *J Am Acad Child Adolesc Psychiatry*. 2021. doi:10.1016/j.jaac.2021.11.031
3. Marlow M, Servili C, Tomlinson M. A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low- and middle-income countries. *Autism Res*. 2019;12(2):176–199. doi:10.1002/aur.2033
4. Sibley MH, Arnold LE, Swanson JM, et al. Variable patterns of remission from ADHD in the multimodal treatment study of ADHD. *Am J Psychiatry*. 2022;179(2):142–151. doi:10.1176/appi.ajp.2021.21010032
5. Riglin L, Wootton RE, Thapar AK, et al. Variable emergence of autism spectrum disorder symptoms from childhood to early adulthood. *Am J Psychiatry*. 2021;178(8):752–760. doi:10.1176/appi.ajp.2020.20071119
6. Landgren M, Kjellman B, Gillberg C. A school for all kinds of minds. *Eur Child Adolesc Psychiatry*. 2003;12(4):162–171. doi:10.1007/s00787-003-0336-0
7. Bahmanyar S, Sundström A, Kaijser M, von Knorring AL, Kieler H. Pharmacological treatment and demographic characteristics of pediatric patients with attention deficit hyperactivity disorder, Sweden. *Eur Neuropsychopharmacol*. 2013;23(12):1732–1738. doi:10.1016/j.euroneuro.2013.07.009
8. Dalsgaard S, Thorsteinsson E, Trabjerg BB, et al. Incidence rates and cumulative incidences of the full spectrum of diagnosed mental disorders in childhood and adolescence. *JAMA Psychiatry*. 2020;77(2):155–164. doi:10.1001/jamapsychiatry.2019.3523
9. Marinopoulou M, Billstedt E, Lin PI, Hallerbäck M, Bornehag CG. Number of words at age 2.5 years is associated with intellectual functioning at age 7 years in the SELMA study. *Acta Paediatr*. 2021;110(7):2134–2141. doi:10.1111/apa.15835
10. Stevanovic D, Knez R, Zorcec T, et al. ESSENCE-Q: slavic language versions for developmental screening in young children. *Neuropsychiatr Dis Treat*. 2018;14:2141–2148. doi:10.2147/NDT.S171359
11. Hatakenaka Y, Maeda M, Ninomiya H, Hachiya K, Fernell E, Gillberg C. ESSENCE-Q obtained in routine Japanese public child health check-ups may be a valuable tool in neurodevelopmental screening. *Acta Paediatr*. 2020;109(4):764–773. doi:10.1111/apa.15029
12. Lindblad I, Westerlund J, Gillberg C, Fernell E. Do all children in primary school have the qualifications to cope with the requirements of the new curriculum? Har alla barn i grundskolan förutsättningar att klara nya läroplanens krav? *Läkartidningen*. 2018;115:EY3L.
13. Busner J, Targum SD. The clinical global impressions scale. *Psychiatry Edmont*. 2007;4(7):28–37.
14. Roid GH, Miller L, Pomplun M, Koch C. Leiter international performance scale – 3rd edition. Stoelting; 2013. Available from: <https://hogrefe.se/klinisk-psykologi/leiter-3/>. Accessed January 3, 2022.
15. Smedje H, Broman JE, Hetta J, von Knorring AL. Psychometric properties of a Swedish version of the “strengths and difficulties questionnaire.”. *Eur Child Adolesc Psychiatry*. 1999;8(2):63–70. doi:10.1007/s007870050086
16. Skolverket KH Resultat på nationella prov i årskurs 3, 6 och 9, läsåret; 2017:27. Available from: <https://www.skolverket.se/download/18.6bfaca41169863e6a65d616/1553968140790/pdf4007.pdf>. Accessed August 31, 2022.
17. The National Agency for Education. Available from: <https://www.skolverket.se/skolutveckling/statistik/fluor-statistiknyheter/statistik/2019-11-28-provresultat-i-grundskolan-lasaret-2018-2019>. Accessed January 2, 2022.
18. Gillberg C. The ESSENCE in child psychiatry: early symptomatic syndromes eliciting neurodevelopmental clinical examinations. *Res Dev Disabil*. 2010;31(6):1543–1551. doi:10.1016/j.ridd.2010.06.002

19. Landgren M, Pettersson R, Kjellman B, Gillberg C. ADHD, damp and other neurodevelopmental/psychiatric disorders in 6-year-old children: epidemiology and co-morbidity. *Dev Med Child Neurol.* 1996;38(10):891–906. doi:10.1111/j.1469-8749.1996.tb15046.x
20. Kadesjö B Neuropsychiatric and neurodevelopmental disorders in a young school-age population. Epidemiology and comorbidity in a school health perspective; 2000. Available from: <https://gupea.ub.gu.se/handle/2077/12134>. Accessed May 6, 2017.
21. Swedish National Agency for Education. Lista över skolors socioekonomiska index 2021.pdf; 2020. Available from: <https://www.skolverket.se/download/18.70f8d1a017495c3cb5913b0/1603700418873/Lista%20%C3%B6ver%20skolors%20socioekonomiska%20index%202021.pdf>. Accessed February 4, 2022.
22. Avvisati F. The measure of socio-economic status in PISA: a review and some suggested improvements. *Large-Scale Assess Educ.* 2020;8(1):8. doi:10.1186/s40536-020-00086-x
23. Taylor MJ, Rosenqvist MA, Larsson H, et al. Etiology of autism spectrum disorders and autistic traits over time. *JAMA Psychiatry.* 2020;77(9):936–943. doi:10.1001/jamapsychiatry.2020.0680
24. Arvidsson O, Gillberg C, Lichtenstein P, Lundström S. Secular changes in the symptom level of clinically diagnosed autism. *J Child Psychol Psychiatry.* 2018;59(7):744–751. doi:10.1111/jcpp.12864
25. Halldner L, Tillander A, Lundholm C, et al. Relative immaturity and ADHD: findings from nationwide registers, parent- and self-reports. *J Child Psychol Psychiatry.* 2014;55(8):897–904. doi:10.1111/jcpp.12229
26. The National Board of Health and Welfare, Socialstyrelsen [Swedish]. Geographic differences in the prescribing of ADHD drugs for children. Geografiska skillnader i förskrivningen av adhd-läkemedel till barn; 2019:33.
27. Russell G, Ford T, Rosenberg R, Kelly S. The association of attention deficit hyperactivity disorder with socioeconomic disadvantage: alternative explanations and evidence. *J Child Psychol Psychiatry.* 2014;55(5):436–445. doi:10.1111/jcpp.12170
28. Björnsdóttir A, Enebrink P, Ghaderi A. Psychometric properties of online administered parental strengths and difficulties questionnaire (SDQ), and normative data based on combined online and paper-and-pencil administration. *Child Adolesc Psychiatry Ment Health.* 2013;7(1):40. doi:10.1186/1753-2000-7-40
29. Cheng S, Keyes KM, Bitfoi A, et al. Understanding parent-teacher agreement of the Strengths and Difficulties Questionnaire (SDQ): comparison across seven European countries. *Int J Methods Psychiatr Res.* 2018;27(1):e1589. doi:10.1002/mpr.1589
30. Ek U, Holmberg K, de Geer L, Swärd C, Fernell E. Behavioural and learning problems in schoolchildren related to cognitive test data. *Acta Paediatr.* 2004;93(7):976–981. doi:10.1111/j.1651-2227.2004.tb02698.x
31. Holmberg K, Hjern A. Health complaints in children with attention-deficit/hyperactivity disorder. *Acta Paediatr.* 2006;95(6):664–670. doi:10.1080/08035250600717121
32. Rydell M, Lundström S, Gillberg C, Lichtenstein P, Larsson H. Has the attention deficit hyperactivity disorder phenotype become more common in children between 2004 and 2014? Trends over 10 years from a Swedish general population sample. *J Child Psychol Psychiatry.* 2018;59(8):863–871. doi:10.1111/jcpp.12882
33. Hatakenaka Y, Fernell E, Sakaguchi M, Ninomiya H, Fukunaga I, Gillberg C. ESSENCE-Q - A first clinical validation study of a new screening questionnaire for young children with suspected neurodevelopmental problems in south Japan. *Neuropsychiatr Dis Treat.* 2016;12:1739–1746. doi:10.2147/NDT.S108411
34. Hatakenaka Y, Ninomiya H, Billstedt E, Fernell E, Gillberg C. ESSENCE-Q- used as a screening tool for neurodevelopmental problems in public health checkups for young children in south Japan. *Neuropsychiatr Dis Treat.* 2017;13:1271–1280. doi:10.2147/NDT.S132546
35. Kattimani S, Joseph KR, Mondal N, Johnson A. Predictive validity of ESSENCE Q screening tool for early detection neurodevelopmental disorder in children. *Asian J Psychiatry.* 2022;73:103162. doi:10.1016/j.ajp.2022.103162
36. Reilly C, Atkinson P, Memon A, et al. Autism, ADHD and parent-reported behavioural difficulties in young children with epilepsy. *Seizure.* 2019;71:233–239. doi:10.1016/j.seizure.2019.08.003

Neuropsychiatric Disease and Treatment

Dovepress

Publish your work in this journal

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the 'PsycINFO' database and CAS, and is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/neuropsychiatric-disease-and-treatment-journal>