Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain

Anna Goodman[1], Robert Goodman[2]

[1] London School of Hygiene & Tropical Medicine, London UK
[2] Kings College London Institute of Psychiatry, London UK

Corresponding author: Anna Goodman, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, UK (anna.goodman@lshtm.ac.uk).

Note: this is a personal version, created by Anna Goodman, of the text of the accepted journal article. It reflects all changes made in the peer review process, but does not incorporate any minor modifications made at the proof stage. The complete citation for the final journal article is:

- *Goodman, A; Goodman, R; (2011) Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain. J Child Psychol Psychiatry. 52 (1) pp. 100–108*
- *DOI: 10.1111/j.1469-7610.2010.02278.x*

Abstract

Background: For adult physical and mental health, the population mean predicts the proportion of individuals with 'high' scores. This has not previously been investigated for child mental health. It is also unclear how far symptom scores on brief questionnaires provide an unbiased method of comparing children with different individual, family or social characteristics.

Methods: Subjects were 18 415 British children aged 5-16 years. Parents, teachers and children aged 11-16 completed Strengths and Difficulties Questionnaires (SDQs) and diagnostic interviews; the latter were used to assign multi-informant clinician-rated diagnoses. We examined how closely the prevalence of child mental disorder was predicted by SDQ mean total difficulty scores, and how these mean scores compared to alternative SDQ-based summary statistics. We did this for populations defined in terms of a wide range of risk factors using one randomly selected half of the study sample. Using these results we generated SDQ prevalence estimator equations, and validated these on the second half of the study sample.

Results: Mean symptom scores closely predicted the prevalence of clinician-rated child mental disorder (R^2=0.89-0.95) and performed better than alternative summary statistics based on binary SDQ outcomes. The predictions of the SDQ prevalence estimators were on average only 1-2% different from the true prevalence, with no systematic tendency towards under- or overestimation. There were only a few outlier subpopulations, all relating to children with learning difficulties.

Conclusion: The proportion of children with a disorder is closely predicted by mean symptom scores, highlighting the potential importance of population-wide interventions to improve child mental health. In Britain, SDQ mean total difficulty scores generally provide an accurate and unbiased method of assessing the mental health of different subgroups. SDQ prevalence estimators may facilitate presenting these research findings as proportions that are more easily interpreted by policymakers and service providers.

Keywords: Strengths and Difficulties Questionnaire; population; mean; prevalence; "SDQ prevalence estimator". **Abbreviations:** DAWBA=Development and Well-Being Assessment; SDQ=Strengths and Difficulties Questionnaire.

KEY POINTS

- For child mental health as for adult mental and physical health, a population's mean symptom score predicts the prevalence of disorder. This highlights the potential importance of research into the determinants of average levels of mental health and of population-wide child mental health interventions.
- The Strengths and Difficulties Questionnaires (SDQ) generally provide accurate and unbiased assessments of child mental health across subgroups defined in terms of multiple child, family and area characteristics.
- SDQ prevalence estimators convert mean scores to approximate prevalences, making them more easily interpreted by policymakers and service providers.

Introduction

Two decades ago, Rose and Day [1] published a seminal demonstration that the population mean of continuously-measured health variables was closely associated with the number of high-scoring ('deviant') individuals. They used this to challenge policymakers and public health practitioners to broaden their focus beyond measures targeted at the high-scoring minority and instead to give greater prominence to population-wide approaches. They likewise called on researchers to devote greater attention to the determinants of the mean scores of populations, and argued that population means may be a valid method for comparing health across groups or monitoring trends over time.

The connection between average population health and the prevalence of disorder has only recently been examined in psychiatric epidemiology. In adults, Veerman et al. [2] used data from five European countries to show that the prevalence of depression was strongly predicted by the population mean on a brief screening questionnaire, the Beck Depression Inventory [3]. For child mental health we know of no work which has addressed this question explicitly. Nevertheless it is widely recognised that emotional, behavioural and hyperactivity disorders in children describe an extreme negative position on a spectrum of symptoms which extends across the full range see e.g. [4 on hyperactivity]. Furthermore, we have previously demonstrated that for the widely-used Strengths and Difficulties Questionnaire SDQ: [5], children with higher total difficulty scores have successively higher probabilities of clinical disorder [6]. This is true for each one-point increase in total difficulty score across the full range, and is seen for the parent, teacher and youth SDQs alike.

These findings imply that differences in the mean total difficulty score in a given population can legitimately be assumed to reflect differences in the prevalence of disorder. In itself, however, this can only tell one about relative differences in prevalence – for example, that disorder prevalence is lower in the intervention arm than in the control arm of a trial. Quantifying differences in mean scores in terms of estimated *absolute* differences in prevalence would often be useful when reporting epidemiological findings for a wider audience. Clinicians, policymakers and the public are likely to find estimates of disorder prevalence more readily interpretable than mean scores on a questionnaire. Service providers would also be able to use the findings directly when planning specialist services.

It would likewise be useful to use mean total difficulty scores to compare different subgroups of children, and then to convert these means into approximate prevalences when disseminating findings. It is, however, unclear whether mean SDQ scores provide an unbiased method of making such comparisons across all types of subgroup. We have previously demonstrated that in some circumstances a given mean total difficulty score on the parent and teacher SDQ had different implications for disorder prevalence in Norway compared to in Britain. Specifically, compared to their British counterparts, Norwegian parents and teachers systematically underreported emotional symptoms on the SDQ [7]. Such reporting biases are also plausible for within-country comparisons. For example, if depressed mothers tend to make unduly negative assessments of their

children's mental health, then a score of (say) 12 points on the parent SDQ might correspond to a prevalence of 20% disorder in children of non-depressed mothers but only 15% in children of depressed mothers. Systematic differences in how informants rate child mental health might also occur with respect to many other potential risk factors, such as socio-economic position or ethnicity. Such systematic differences would undermine the validity of using mean scores to compare the mental health of subgroups, and would imply that no single conversion algorithm could be used to convert SDQ mean scores into prevalence estimates.

Moreover, even if mean total difficulty score *did* prove an unbiased predictor of prevalence across all subgroups, this would not imply that it was always the best method available. For example, using the supplementary SDQ questions on impact and triangulating symptom and impact scores across informants improves the identification of disorders in individual children [8, 9]. It is plausible that summary statistics incorporating this additional information would likewise measure population mental health more accurately and thereby improve the estimation of disorder prevalence. Alternatively, for audiences such as policymakers, it might sometimes be appropriate to focus on a highly transparent, informant-centred measure. The SDQ question that asks respondents whether the young person has a definite or severe mental health problem is simple and transparent, but it is unknown to what extent this provides an unbiased method of making comparisons across subgroups.

In this paper we therefore use a large, nationally-representative samples of British 5-16 year olds to: 1) present the first formal investigation of whether a population's mean child mental health symptom scores (measured using the SDQ) predicts the prevalence of disorder; 2) examine how far all subgroups show the same relationship between SDQ mean scores and population prevalence; 3) compare the performance of mean SDQ symptom score with alternative SDQ-based measures; and 4) validate 'SDQ prevalence estimators' converting mean SDQ total difficulty scores into approximate prevalence estimates.

Methods

Sample

Our data come from the British Child and Adolescent Mental Health Surveys of 1999 and 2004. These were two nationally-representative, population-based surveys of the mental health of children and adolescents (henceforth "children") aged 5-16 in Great Britain for full details, see [10, 11]. Primary caregivers ('parents') were approached to give informed consent for face-to-face interview. With parental permission, teachers and young people aged 11-16 were also approached. All respondents were asked to complete the Strengths and Difficulties Questionnaire (SDQ), a diagnostic interview and information on potential risk factors. Diagnostic interview data from all available respondents were used to assign clinical diagnoses, which were available for a total of 18 415 children (mean age 10.2 years, 50.8% male). Of these, complete SDQ data was available for 18 130 parents, 13 990 teachers and 7483 11-16 year olds.

The British Child and Adolescent Mental Health Surveys received ethical approval from the Research Ethics Committee of the Institute of Psychiatry, King's College London, and the national Multi-Centre Research Ethics Committee for England, Scotland and Wales.

Measures

Strengths and Difficulties Questionnaire (SDQ)

The Strengths and Difficulties Questionnaire (SDQ) is a measure of mental health problems in children aged 4-17 which can be administered to parents, teachers and young people aged 11 or over [5, 12, 13]. Its 20 items relating to emotional symptoms, conduct problems, hyperactivity and peer problems are summed to create a 'total difficulty score' ranging from 0-40. These symptom questions are then followed by a single item which asks whether the respondent thinks that the child or adolescent has a problem with "emotions, concentration, behaviour or being able to get along with other people" (response options: No problems, Minor problems, Definite problems and Severe problems). Finally, the SDQ impact supplement asks about how much distress or impairment is caused by any problems identified. We identified children with borderline or high symptoms plus high impact ('symptom+impact') from single informants [14]. We also triangulated information across informants to generate a multi-informant predictor, using a previously-validated algorithm which mimics operationalised diagnostic criteria by requiring both symptoms and impact, and by specifying the need for pervasiveness in the case of hyperactivity [8, 9].

In this paper we report data on ten potential methods of using the SDQ to predict population prevalence. For the parent, teacher and youth SDQ we calculate 1) mean total difficulty score, 2) the proportion of individuals with symptoms+impact and 3) the proportion of individuals reporting 'definite' or 'severe' problems. Our tenth measure is the proportion of children with symptoms+impact as defined by triangulating across multiple informants.

Development and Well-being Assessment (DAWBA)

Our measure of disorder prevalence comes from the Development and Well-being Assessment (DAWBA). This is a detailed psychiatric interview for parents and 11-16 year olds, and a briefer questionnaire for teachers [15]. Each section of the DAWBA uses screening questions, but children are also administered the section in full if they have a high score on the relevant SDQ subscale (for example, the hyperactivity SDQ subscale for the hyperactivity disorder section). Each section contains fully-structured questions followed by open-ended descriptions by respondents of any problem areas. Experienced clinicians then use both the closed and open DAWBA responses, and triangulate information from across informants, in order to assign diagnoses according to DSM-IV [16]. These diagnoses have been shown to have good reliability and validity in British samples [15, 17].

Analyses

We randomly split our study sample into two halves. Using the first half we examined the relationship between population prevalence and each of our ten SDQ estimators of interest. We did this for populations defined by stratifying the total sample according to the seven risk factors shown in the first Column of Table 1, chosen *a priori* to span from very high-risk to very low-risk groups. The number of individuals in each risk factor strata is presented in the Electronic appendix, and was in all cases at least 250 parent SDQs (200 teacher SDQs, 90 youth SDQS). Each risk factor populations was then given equal weight when fitting linear regression models. We used the logodds of the population prevalence as our outcome when examining the mean total difficulty scores because we have previously demonstrated a sigmoid relationship between total difficulties scores and the prevalence of disorder [6]. For the proportion with SDQ symptoms+impact and the proportion with definite/severe problems, we used the untransformed population prevalence. We fitted separate regression models for each of our ten SDQ predictors, with the outcome being the population prevalence in the same subset of children (e.g. comparing predictors based on teacher SDQs with the prevalence of disorder in children with teacher SDQ data). We used the adjusted R^2 of each model as a measure of the variance in the prevalence explained.

We initially fitted these regression models after pooling together both genders and children of all ages. To fine-tune the algorithm for converting SDQ mean scores into prevalence estimates, we then stratified all populations by gender and by age (5-10 years vs. 11-16 years). We tested whether a population's gender or age showed an interaction with SDQ score and/or a main effect when predicting prevalence, and incorporated any effects into the prevalence estimator formula. We singled out age and gender for this treatment because almost all studies collect this data and many also build these characteristics into their sampling frame (e.g. sampling from a girls' secondary school). We then validated the resulting SDQ prevalence estimator equations on the second half of the sample. We did this by comparing predicted vs. observed prevalence for populations defined according the seven risk factors used to derive the equations and also for a further eleven risk factors (Table 1, Column 2)

Table 1: Child, family, school and family risk factors used to derive and to validate the British SDQ prevalence estimator equations

	Column 1: Variables used to derived the SDQ prevalence estimators	Column 2: Additional variables used to validate the SDQ prevalence estimators
Area and school	[I] Small area deprivation, from the most recent Indices of Multiple Deprivation 2004 for England and Scotland, 2005 for Wales: [18]: using deciles.	[I] Ford Score, a predictor of the prevalence of mental health problems in a school [19]: 0-2 points, 3-5 points, 6-8 points, 9-11 points, 12-17 points.
Socio-economic position	[P] Responding parent's highest educational qualification: no qualifications; poor GCSEs (grades D-F) or equivalent; good GCSEs (grades A-C); A-level; diploma; degree.	[P] Weekly household income: £0-99; £100-199; £200-299; £300-399; £400-499; £500-599; £600-769; £770 or over. [P] Housing tenure: owner occupied, social sector rented, privately rented. [P] Occupational social class: I, II, III Non-manual, III Manual, IV, V, Never worked/full-time student.
Family composition	[P] Family type: two-parent, stepfamily, lone parent family	[P] Mother's age at child's birth: 19 years or under, 20-24 years, 25-29 years, 30-34 years, 35-39 years, 40 years or more.
Family stress	[P*] Parent's mental health, from the 12-item General Health Questionnaire GHQ-12: [20]: 0-1 points, 2-3 points, 4-5 points, 6-7 points, 8-9 points, 10-12 points.	[P*] Family functioning, from the general functioning subscale of the McMaster Family Activity Device [21]: 1-1.49 points, 1.5-1.99 points, 2-2.49 points, 2.5-3.99 points.
Child	[P] General physical health: very good, good, very bad/bad/fair. [P] Parent-reported learning difficulties or dyslexia: yes, no	[P] Ethnicity: White, Black, Indian, Pakistani/Bangladeshi, Other. [T] Teacher-reported academic abilities, created by summing the teacher's response to three questions on the child's ability in maths, reading spelling: 0-1 points (approx. 0-5th percentile), 2-3 points (5-20th percentile), 4-5 points (20-35th percentile), 6-7 points (35-70th percentile), 8-9 points (70-100th percentile). [P] Number of stressful life events ever experienced (parental separation , financial crisis, family police contact, death of parent/sibling, serious illness, serious accident, death of a friend): 0 events, 1 event, 2 events, 3 events or more [P] Neurodevelopmental disorder (cerebral palsy, epilepsy, muscle disease or weakness, or co-ordination problems) with no parent-reported learning difficulty: yes, no [P] Neurodevelopmental disorder and a parent-reported learning difficulty: yes, no
Combined category	"Supernormal" = Lowest-risk category on all the Column 1 variables: yes, no.	

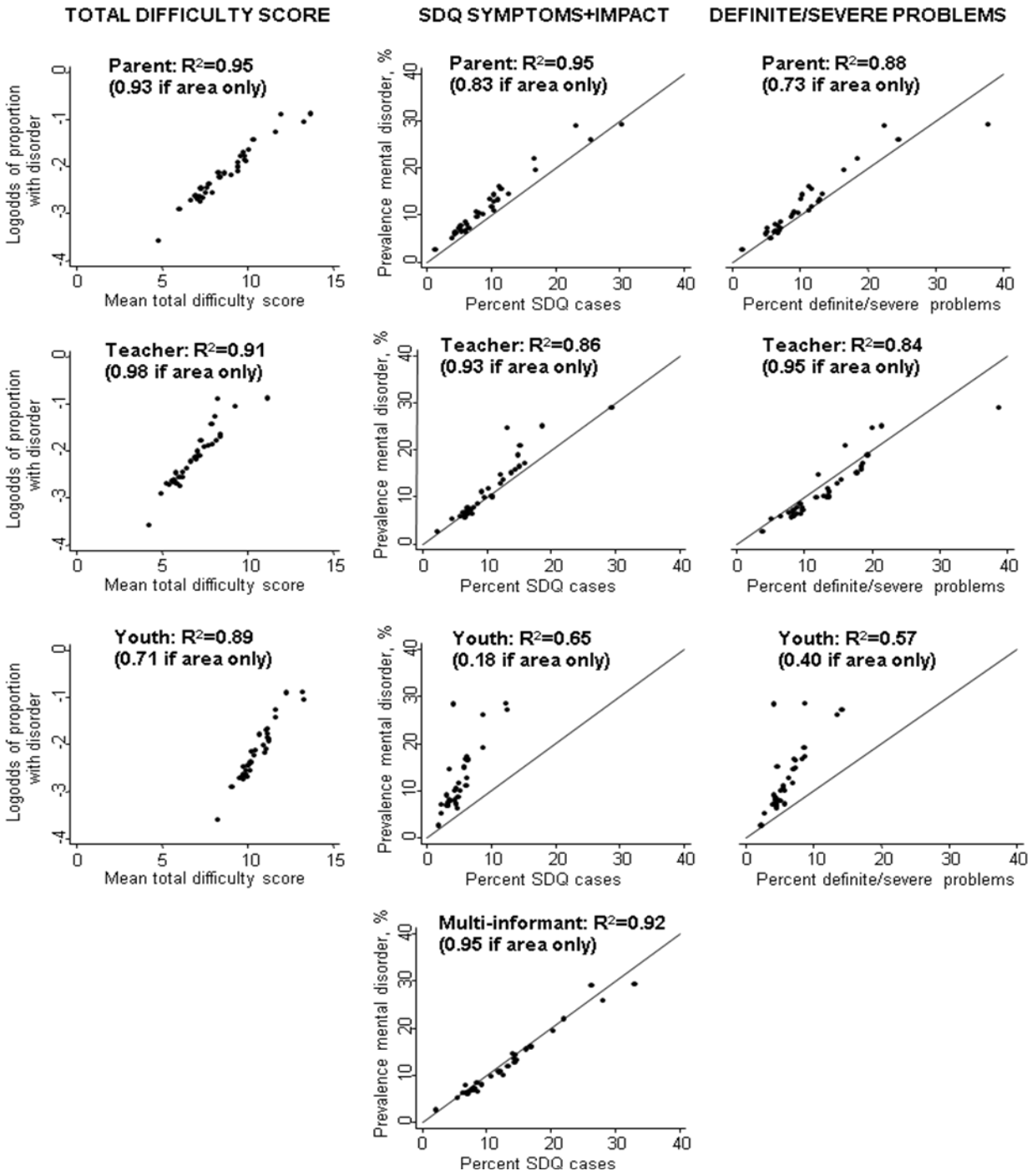
[P]=parent-reported; [T]=teacher reported; [C]=child-reported; [I]=assigned by investigators. *=sensitive items administered by laptop.

Results

Relationship between mean total difficulty score and population prevalence

Mean total difficulty scores for the parent, teacher and youth SDQs all showed a linear relationship with the logodds of the population prevalence, and this relationship applied to subgroups defined in terms of all selected child, family and area risk factors. All three mean scores were highly predictive of disorder prevalence, with adjusted R^2 values of 0.89-0.95 – i.e. explaining 89-95% of the observed variance (Figure 1, first column). These values showed only a modest decrease after excluding children with disorders from the SDQ mean scores (R^2 values 0.76-0.88). These values were very similar for parent and teacher SDQs in analyses restricted to the ten non-overlapping populations defined by deciles of small-area deprivation (R^2=0.93 for parents and 0.98 for teacher), although somewhat lower for the youth SDQ (R^2=0.71).

Figure 1: Association between SDQ outcomes and population prevalence.



Analyses were based on 32 populations defined in terms of the variables in the first column of Table 1, using first randomly-selected half of the study sample (N=9036 parents, 6955 teachers, 3765 young people). The diagonal lines represent perfect agreement between the percentages obtained from the 'symptoms+impact 'and 'definite/severe problems' SDQ measures and the prevalence of disorder. 'Area only' analyses use only the ten populations defined in terms of deciles of small-area deprivation.

Comparing mean total difficulty score with other SDQ-based summary statistics

The R^2 values for the mean total difficulty scores were always at least as high as the corresponding SDQ symptoms+impact measure (Figure 1, second column) and were also comparable to the multi-informant symptoms+impact predictor (R^2=0.92). Thus at the population-level, the potential measurement gain from incorporating impact information and triangulating across informants seemed to be more than offset by greater measurement error due to using a binary summary statistic. Nevertheless, high R^2 values (0.88-0.95) were also observed on the symptoms+impact predictors from the parent, teacher and multi-informant SDQs (but not the youth SDQ). This indicates that these are legitimate alternatives for comparing mental health across British subgroups if the total difficulty score mean is not available (e.g. in meta-analyses of published information). Moreover, both the teacher and multi-informant predictors generated proportions with very similar absolute values to the population prevalence of disorder. By contrast, prevalence was systematically underestimated by the parent SDQ and even more markedly underestimated by the youth SDQ.

Finally, the single-item assessment of whether the child had definite/severe mental health problems performed surprisingly well when made by parents or teachers (R^2 values 0.84-0.88). This is inferior to the performance of the mean total difficulty score, and these are therefore not the optimal methods of comparing subgroups of children. Nonetheless, they do appear a viable alternative to more complicated SDQ-based outcomes in situations in which transparency and simplicity is felt to be of particular importance.

Additionally adjusting for age and gender to derive the British SDQ prevalence estimators

The graphs presented in Figure 1 present populations which combine boys with girls and which combine 5-10 year olds with 11-16 year olds. When repeated after stratifying by these characteristics, there was never any evidence ($p>0.05$) of an interaction between total difficulty score and either age or gender in predicting population prevalence. By contrast, there was evidence ($p\leq0.001$) that age had a main effect upon prevalence when using the parent and teacher SDQ; and that gender had a main effect when using the teacher and youth SDQ (Figures stratifying the parent, teacher and child SDQ by age and gender are presented in the Electronic appendix). The SDQ prevalence estimators presented in

Table 2 therefore incorporate these main effects in order to allow the equations to be applied to samples with age and gender compositions different to that of the British Child and Adolescent Mental Health Surveys. Note that the equation within the square brackets corresponds to the linear regression line plotting each population against the logodds of the prevalence, using the same populations as in Figure 1 but additionally stratifying by age and gender. For look-up graphs plotting mean SDQ scores against population prevalence, see the Electronic Appendix and also www.sdqinfo.com/prevalence_estimators.

Table 2: British SDQ prevalence estimators for the parent, teacher and youth SDQs

	Proportion of individuals with a disorder: for percent, multiply by 100
Parent	1 /(1 + exp(- [(0.303*TDS) + (0.539* p(Old)) - 5.102]))
Teacher	1 /(1 + exp(- [(0.394*TDS) + (0.450* p(Old)) + (0.411* p(Fem)) - 5.313]))
Youth	1 / (1 + exp(- [(0.524*TDS) - (0.220* p(Fem)) - 7.419]))

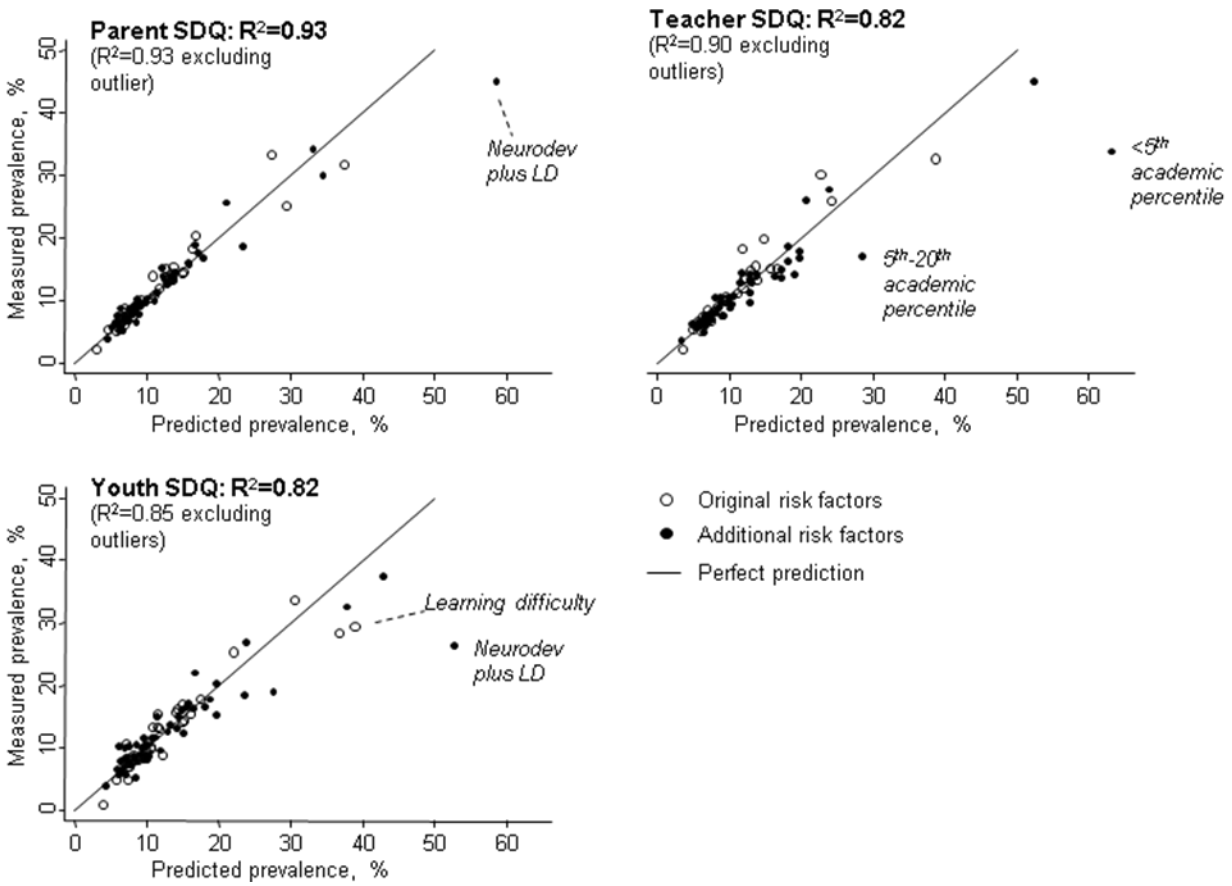
exp=exponential; TDS=total difficulty score; p(Old)=proportion of sample aged 11-16 (vs. 5-10); p(Fem)=proportion of sample female.

Validating the British SDQ prevalence estimators

We validated the British SDQ prevalence estimators shown in

Table 2 using the second randomly-selected half of our data sample. In general the prevalence estimators performed well (R^2 0.82-0.93), and this good performance extended to the eleven additional risk factors which were not used when deriving the estimator equations (see **Error! Reference source not found.**). For all three informants, the mean absolute difference between the predicted and the measured prevalence was only 1-2% and in no case was there evidence ($p<0.05$) of any systematic tendency towards over- or underestimation (Table 3). For all three classes of informant, however, there were one or two outlier subpopulations in which the predicted prevalence from the SDQ score overestimated the true prevalence by more than 10 percentage points. These outliers were all subpopulations with learning difficulties or low academic abilities, suggesting that mean SDQ scores may generate misleadingly high prevalence estimates for this particular group.

Figure 2: Prevalence of child mental disorder predicted by the British SDQ prevalence estimators as compared to the measured prevalence



Neurodev plus LD=neurodevelopmental disorder plus learning difficulties. Original risk factors = 32 populations defined using variables in Table 1, Column 1; Additional risk factors = 51 additional populations defined using variables in Table 1, Column 2. Analyses were based on the second randomly-selected half of the study sample (N=9094 parents, 7035 teachers, 3718 young people). The diagonal lines represent perfect agreement between the percentages obtained from the SDQ prevalence estimators and the measured prevalence of disorder. Outlier subpopulations are labelled, being defined as populations with an absolute discrepancy of more than 10 percentage points between the predicted and the measured prevalence

Table 3: Validating the British SDQ prevalence estimators; mean absolute and systematic discrepancy between the predicted and the measured prevalences.

	Mean absolute discrepancy & 95% confidence interval	**Mean systematic discrepancy (bias) & 95% confidence interval**
Parent SDQ	1.3% (0.9%, 1.7%)	-0.2% (-0.7%, 0.3%)
Teacher SDQ	2.0% (1.2%, 2.8%)	0.2% (-0.7%, 1.1%)
Youth SDQ	2.1% (1.4%, 2.8%)	0.4% (-0.4%, 1.3%)

Absolute discrepancy was calculated as the difference between the predicted prevalence and the measured prevalence without regard to the sign of the difference. Systematic discrepancy (bias) was calculated as the predicted prevalence minus the measured prevalence, with regard to the sign of the difference. Analyses were based on the second randomly-selected half of the study sample (N=9094 parents, 7035 teachers, 3718 children).

Discussion

In this representative sample of 18 415 British young people aged 5-16 years, we have demonstrated that a population's mean symptom score closely predicts the prevalence of clinician-rated child mental disorder. This was true for symptom scores reported by parents, teachers and young people alike, as measured using the widely-used Strengths and Difficulties Questionnaire (SDQ). It was also generally true with reference to 'populations' defined in terms of a wide range of child, family, school and area risk factors. Mean symptom scores always performed better at predicting prevalence than alternative population summary statistics based on binary SDQ outcomes. The SDQ prevalence estimators that we created using one half of the sample performed well when applied to the SDQs collected in the remaining half of the sample; the estimates they generated were on average only 1-2% different from the true prevalence, with no systematic tendency towards under- or overestimation. There were only a few outlier subpopulations, all relating to children with learning difficulties. We conclude that SDQ mean total difficulty scores from any informant generally provide an accurate and unbiased method for monitoring or comparing the mental health of different subgroups of British children; and that the SDQ prevalence estimators represent a potentially useful tool for presenting research findings for a wider audience.

Before considering the theoretical and practical importance of these findings, it is worth highlighting some this study's strengths and limitations. One important strength was the administration of questionnaire and diagnostic interview measures to all children. Unlike previous two-phase studies in adults [2], we therefore did not need to impute disorder status based upon questionnaire score. Nevertheless, a small potential for circularity remains because receiving high SDQ scores leads to some DAWBA sections being administered in full even if respondents do not screen positive on the DAWBA's own screening questions. Collecting this additional DAWBA information is occasionally the basis for assigning diagnoses which would otherwise have been missed. This cannot explain the results observed, however, as a strong association with prevalence remained after excluding the mean scores of children with a disorder.

A more serious limitation is that our study drew exclusively on British data so the resultant SDQ prevalence estimators cannot be assumed to be valid for non-British samples. For physical health measures recorded using objective measures, there is evidence that the association between the population mean and prevalence is observed globally [1]. Likewise for adult depression, the same association between symptoms and impact was observed within-countries and internationally across five European countries [2]. It is also worth noting that although based on small numbers, the SDQ prevalence estimators seemed to work well in Black, Indian, Pakistani/Bangladeshi and 'Other' ethnic groups. These findings all provide some grounds for optimism regarding the generalisability of the British SDQ prevalence estimators. On the other hand, we have previously shown that Norwegian parents and teachers systematically underreport emotional symptoms on the SDQ as compared to their British counterparts [7]. We are not aware of any other published studies that have used detailed diagnostic measures to investigate the potential for reporting biases using brief child mental health screening

questionnaires; without such studies it cannot be assumed that a given SDQ score means the same thing across different countries.

Yet while this study cannot address the possibility of international reporting biases, it does indicate that *within* Britain the SDQ is an unbiased predictor of mental health across a very wide range of child, family, school and area factors. The only important exception was that children with learning difficulties appear to have received misleadingly high total difficulty scores. Otherwise, these findings indicate that SDQ differences between subgroups of British children can legitimately be interpreted as reflecting genuine mental health differences rather than reporting bias. Mean total difficulty scores provided better prediction than alternative SDQ-based measures in predicting disorder, suggesting that these are the method of choice for researchers seeking to compare and monitor mental health. The parent, teacher and youth SDQ all performed well, and for all three informants we validated a new tool for generating ball-park prevalence estimates based on these mean SDQ scores. For parents and teachers, we also showed relatively good performance by highly-transparent single-item reports of whether a child had 'definite or severe problems'. We believe these findings may have substantial practical value in terms of translating epidemiological findings to a form which UK policymakers can readily interpret and service planners can readily act upon.

Our paper also raises issues of wider theoretical importance. One is the message that the optimal use of brief screening questionnaires may differ when studying populations as opposed to individuals. This is exemplified by the absence of superiority of summary statistics based on SDQ symptoms+impact, despite the fact that their incorporation of impact and/or triangulation across informants improves disorder prediction at the individual level [9]. Another issue of wider theoretical importance is the demonstration that the prevalence of child mental disorder is closely predicted by that population's mean score; in analyses of the ten non-overlapping populations defined by deciles of small-area deprivation, mean symptom scores explained 93% of the variance in prevalence for the parent SDQ, 98% for the teacher SDQ and 71% for the child SDQ. These parent and teacher SDQ values are higher than the figure of 84% reported in a recent cross-national comparison of depression symptoms and depressive disorder in adults [2] and also higher or similar to the values of 61-94% obtained by Rose and Day [1] for blood pressure, overweight, sodium intake and alcohol intake. The association between the population mean and the prevalence of disorder thus appears to be at least as strong for child mental health as for the physical and mental health of adults. Moreover, this strong association could not simply be explained by children with disorders bringing up the population average; rather a substantial correlation remained after excluding children with a disorder from the population mean.

These findings underline that child mental disorders represent the extreme end of a distribution rather than a category which is wholly distinct from the normal range. Moreover, the proportion of children with a disorder is a function of the properties of the distribution as a whole; "the minorities' problems exist as a consequence of the majority's attributes...[and] the health of society is integral" [1, p.1034]. To the extent that this applies to child mental health, researchers should consider investigating the determinants of average

mental health and policymakers should consider implementing population-wide interventions alongside more targeted approaches. By highlighting this and by also providing practical new tools for speaking to policymakers, we hope this paper will contribute to a future in which the mental health of *all* children is taken seriously.

Conflict of interest

AG and RG are directors and RG is the owner of Youthinmind, which provides no-cost and low-cost software and web sites related to the SDQ and the DAWBA.

Acknowledgements

The British Child and Adolescent Mental Health Surveys were carried out by the Office for National Statistics, and paid for by the Department of Health and the Scottish Health Executive.

References

1. Rose G, Day S. The population mean predicts the number of deviant individuals. *BMJ.* Nov 3 1990;301(6759):1031-1034.
2. Veerman JL, Dowrick C, Ayuso-Mateos JL, Dunn G, Barendregt JJ. Population prevalence of depression and mean Beck Depression Inventory score. *Br J Psychiatry.* Dec 2009;195(6):516-519.
3. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry.* Jun 1961;4:561-571.
4. Taylor E, Sandberg S, Thorley G, Giles S. *The Epidemiology of Childhood Hyperactivity, Maudsley Monograph No. 33*. Oxford: Oxford University Press; 1991.
5. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry.* Jul 1997;38(5):581-586.
6. Goodman A, Goodman R. Strengths and difficulties questionnaire as a dimensional measure of child mental health. *J Am Acad Child Adolesc Psychiatry.* Apr 2009;48(4):400-403.
7. Heiervang E, Goodman A, Goodman R. The Nordic advantage in child mental health: separating health differences from reporting style in a cross-cultural comparison of psychopathology. *J Child Psychol Psychiatry.* 2008;49(6):678-685.
8. Goodman R, Renfrew D, Mullick M. Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *Eur Child Adolesc Psychiatry.* Jun 2000;9(2):129-134.

9. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br J Psychiatry.* Dec 2000;177:534-539.

10. Meltzer H, Gatward R, Goodman R, Ford T. *Mental health of children and adolescents in Great Britain*. London: The Stationery Office; 2000.

11. Green H, McGinnity A, Meltzer H, Ford T, Goodman R. *Mental health of children and young people in Great Britain, 2004*: Palgrave MacMillan; 2005.

12. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry.* Nov 2001;40(11):1337-1345.

13. Goodman R, Scott S. Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful? *J Abnorm Child Psychol.* Feb 1999;27(1):17-24.

14. Goodman R, Ford T, Meltzer H. Mental health problems of children in the community: 18 month follow up. *Bmj*. Jun 22 2002;324(7352):1496-1497.

15. Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatry.* Jul 2000;41(5):645-655.

16. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV).* Washington, DC: American Psychiatric Association.; 1994.

17. Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV disorders. *J Am Acad Child Adolesc Psychiatry.* Oct 2003;42(10):1203-1211.

18. Noble M, Wright G, Dibben C, et al. *Indices of Deprivation 2004*. London: Office of the Deputy Prime Minister; 2004.

19. Goodman A, Ford T. Validation of the Ford score as a measure for predicting the level of emotional and behavioural problems in mainstream schools. *Research in Education.* 2008;80(1):1-14.

20. Goldberg DP, Williams P. *A user's guide to the General Health Questionnaire*. Windsor: NFER-Nelson 1998.

21. Miller I, Epstein N, Bishop D, Keitner G. The McMaster Family Assessment Device: reliability and validity. *Journal of Marital and Family Therapy.* 1985;11:345-356.

Electronic Appendix

Table 4: Number of parent, teacher and youth SDQs in each of the populations defined by risk factors and used in regression analyses

	Variable	Level	No. SDQS from		
			Parents	Teachers	Youth
	TOTAL SAMPLE		18 130	13 990	7483
Variables used to derived the SDQ prevalence estimators	**Small area deprivation (Index of Multiple Deprivation)**	Decile 1 (least deprived)	1951	1564	854
		Decile 2	1939	1573	830
		Decile 3	1934	1557	874
		Decile 4	1852	1489	763
		Decile 5	1891	1457	774
		Decile 6	1830	1401	789
		Decile 7	1781	1371	708
		Decile 8	1718	1302	714
		Decile 9	1602	1127	580
		Decile 10 (most deprived)	1617	1136	587
	Parent's highest educational qualification	No qualification	3792	2781	1679
		Poor GCSEs (grades D-F)	2548	1955	1046
		Good GCSEs (grades A-C)	5364	4174	2149
		A-level	1948	1536	753
		Diploma	2066	1627	901
		Degree	2320	1873	944
	Family type	Two-parent	11 869	9361	4774
		Stepfamily	2021	1547	953
		Lone parent family	4203	3054	1740
	Parent's mental health (GHQ-12)	0-1 points	12 119	9399	4869
		2-3 points	2566	2029	1085
		4-5 points	1305	987	574
		6-7 points	860	649	395
		8-9 points	580	442	264
		10-12 points	549	405	263
	General physical health	Very bad/bad/fair	1190	859	469
		Good	4393	3348	1797
		Very good	12 545	9781	5217
	Learning difficulties or dyslexia	No	16 316	12 629	6720
		Yes	1813	1360	763
	"Supernormal" (Lowest--risk on all the above variables	No	17 088	13 139	7098
		Yes	1042	851	385
Additional variables used to validate the SDQ prevalence estimators	**Ford Score**	0-2 points	3149	2659	1317
		3-5 points	4990	4006	2116
		6-8 points	3793	2923	1538
		9-11 points	1619	1246	554
		12-17 points	433	295	261
	Weekly household income	£0-99	709	530	244
		£100-199	2653	1954	988
		£200-299	2393	1768	988
		£300-399	1999	1573	812
		£400-499	1890	1460	819
		£500-599	1667	1322	733
		£600-769	2188	1742	965
		£770 or over	3425	2799	1491

	Variable	Level	No. SDQS from		
			Parents	**Teachers**	**Youth**
	Housing tenure	Owner occupied	12 594	9915	5424
		Social sector rented	4315	3150	1604
		Privately rented.	1215	923	455
	Occupational social class	I	982	812	418
		II	5233	4200	2293
		III Non-manual	3509	2722	1487
		III Manual	3201	2431	1289
		IV	3338	2486	1326
		V	870	636	379
		Never worked/ student.	457	346	142
	Mother's age at child's birth	19 years or under	966	723	382
		20-24 years	3902	2899	1724
		25-29 years	6056	4724	2589
		30-34 years	4540	3557	1729
		35-39 years	1715	1367	619
		40 years or more	315	244	115
	Family functioning (General functioning scale of Family Activity Device)	1-1.49 points	5329	4225	2003
		1.5-1.99 points	7581	5860	3241
		2-2.49 points	4381	3296	1891
		2.5-3.99 points	618	479	276
	Ethnicity	White	16 390	12760	6814
		Black	433	289	175
		Indian	364	259	152
		Pakistani/Bangladeshi	374	263	120
		Other	565	417	221
	Teacher-reported academic abilities	0-1 points	4290	4227	1997
		2-3 points	4932	4870	2018
		4-5 points	1930	1902	576
		6-7 points	1767	1735	648
		8-9 points	857	847	199
	Number of stressful life events	0 events	8236	6453	3138
		1 event	6017	4681	2519
		2 events	2633	1966	1227
		3 events or more	1157	846	585
	Neurodevelopmental disorder without learning difficulty	No	17 859	13 788	7364
		Yes	271	202	119
	Neurodevelopmental disorder with learning difficulty	No	17 813	13 746	7391
		Yes	317	244	92

Note: missing data for some risk factors means that in some cases the numbers of SDQs across the different strata sums to less than the total sample.

Figure 3: Parent, teacher and child SDQ mean total difficulty scores vs. the logodds of the prevalence of disorder, stratified by age and gender.

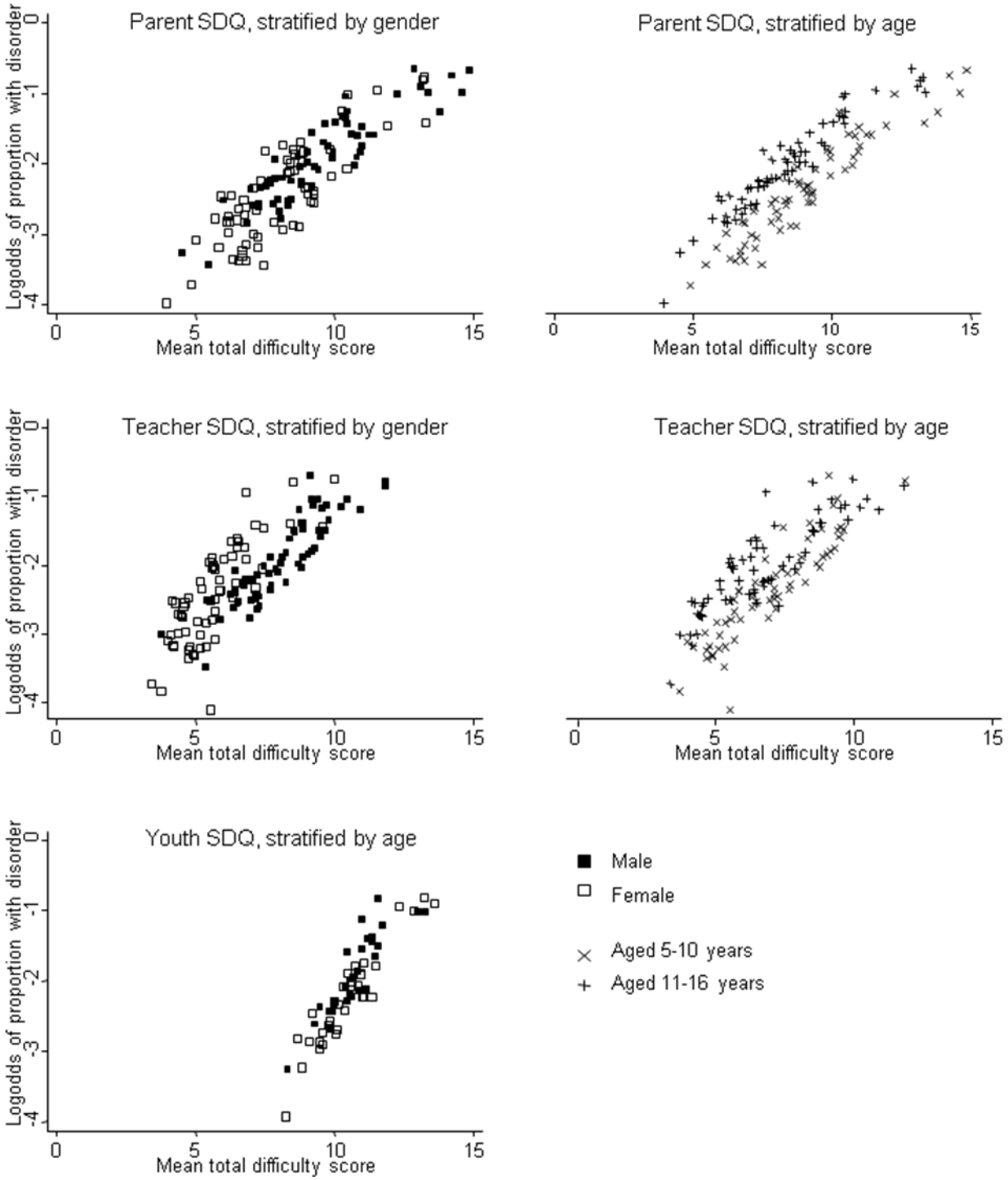


Figure 4: UK norms for the estimated prevalence of disorders, based upon mean total difficulty scores on the parent SDQ.

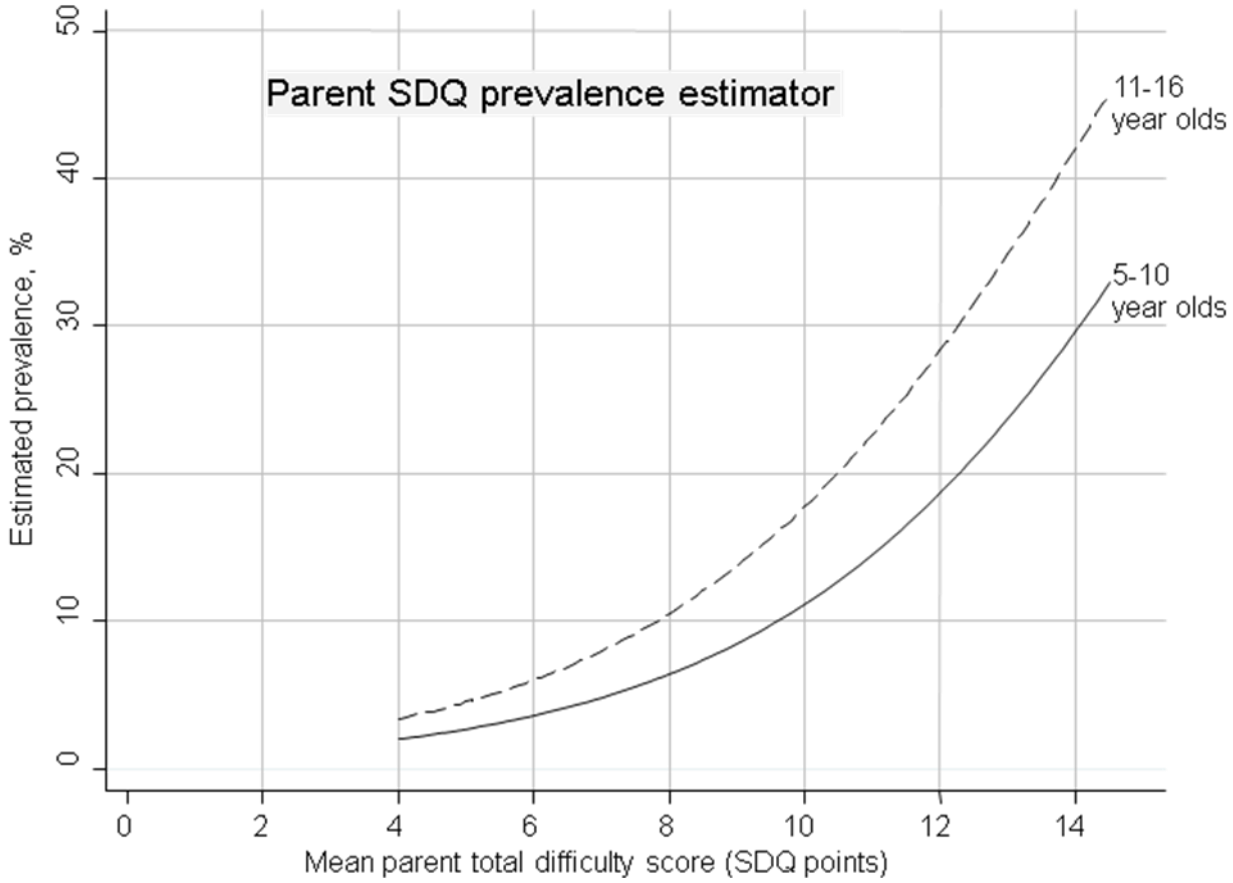


Figure 5: UK norms for the estimated prevalence of disorders, based upon mean total difficulty scores on the teacher SDQ.

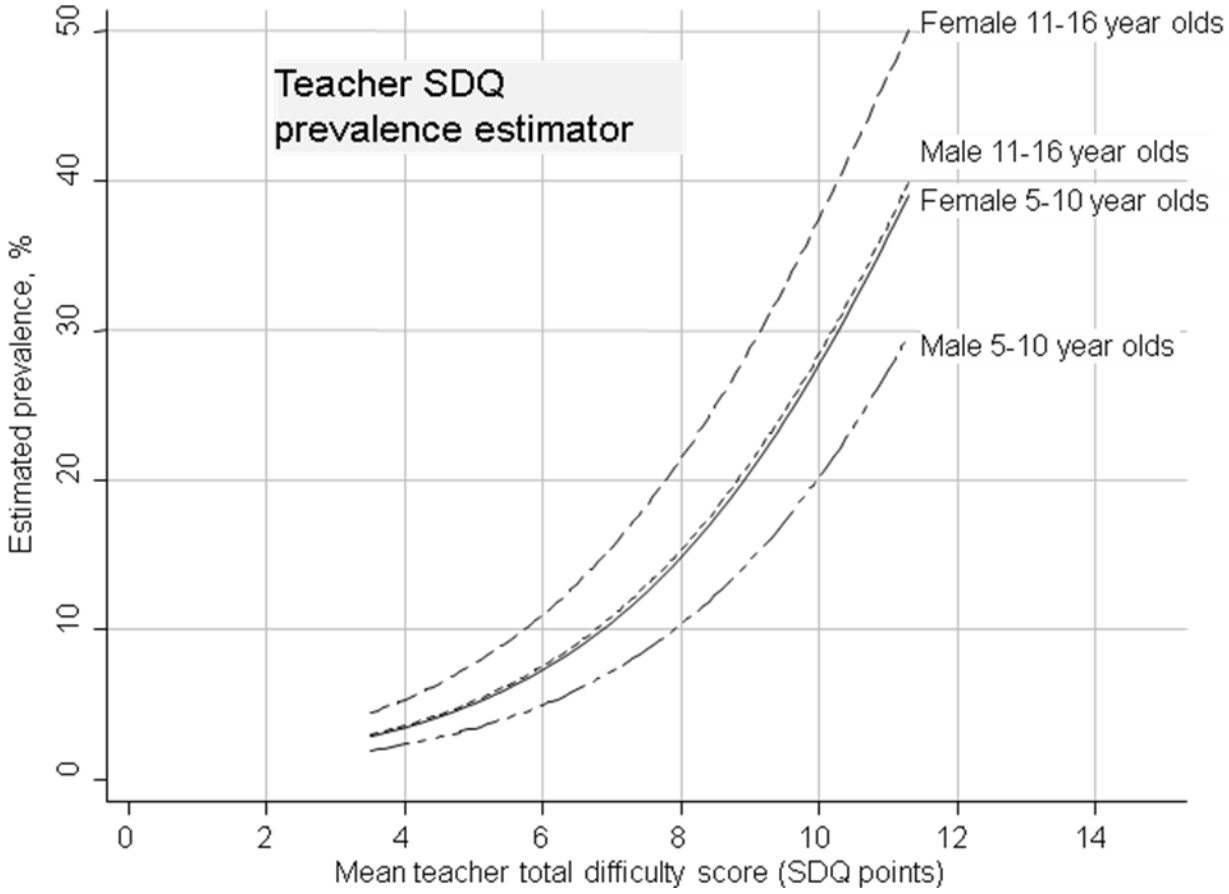


Figure 6: UK norms for the estimated prevalence of disorders, based upon mean total difficulty scores on the youth SDQ