Potential value of the current mental health monitoring of children in state care in England

Christine Cocker, Helen Minnis and Helen Sweeting

Background
Routine screening to identify mental health problems in English looked-after children has been conducted since 2009 using the Strengths and Difficulties Questionnaire (SDQ).

Aims
To investigate the degree to which data collection achieves screening aims (identifying scale of problem, having an impact on mental health) and the potential analytic value of the data set.

Method
Department for Education data (2009–2017) were used to examine: aggregate, population-level trends in SDQ scores in 4/5–16/17-year-olds; representativeness of the SDQ sample; attrition in this sample.

Results
Mean SDQ scores (around 50% ‘abnormal’ or ‘borderline’) were stable over 9 years. Levels of missing data were high (25–30%), as was attrition (28% retained for 4 years). Cross-sectional SDQ samples were not representative and longitudinal samples were biased.

Conclusions
Mental health screening appears justified and the data set has research potential, but the English screening programme falls short because of missing data and inadequate referral routes for those with difficulties.

Declaration of interest
None.

Keywords
Looked-after children; mental health; screening; Strengths and Difficulties Questionnaire.

The mental health of children in state care is of great concern. Because of this, in 2009 the Department for Education introduced compulsory mental health data collection for these children in England by using the Strengths and Difficulties Questionnaire (SDQ). This article examines the degree to which the current mass data collection achieves screening aims and the potential analytic value of the resulting data set.

The SDQ
The SDQ is an internationally validated questionnaire, comprising 25 items, which are broken down into five scales: emotional symptoms, conduct problems, hyperactivity, peer problems, social behaviour. A general difficulties score is created by adding up the scores from the first four scales. The cut-offs for this score were originally chosen ‘so that roughly 80% of children in the community are normal, 10% are borderline, and 10% are abnormal’. There are (almost identical) versions for completion by: parents/carers/teachers of 4–16-year-olds; parents/carers/teachers of 3–4-year-olds and 11–16-year-olds themselves. In addition to high specificity (80%) and sensitivity (85%), the main benefits of the SDQ are that it is free, quick and straightforward to use. However, cross-informant agreement tends to be lower for internalising than for (more observable) externalising behaviours, and emotional symptoms are best identified by self-reports. The SDQ is one of the most used and recognised child and adolescent screening tools. In the UK, it has been successfully used to screen for child psychiatric disorders in both community and looked-after children samples, with the study of looked-after children concluding that ‘screening with the SDQ (carer and teacher versions) could improve the detection and treatment of behavioural, emotional, and concentration problems among looked-after children’.

Data collection for looked-after children in England
In England, it is compulsory to collect mental health data (using the carer-report SDQ) from all children aged 4/5–16/17 who have been in state care for 1 year or longer. The mental health of these children is known to be poor and routine SDQ data is seen by the Department for Education as both a way of identifying the scale of the problem and, at an individual level, of highlighting the likelihood that the child either has, or could develop significant mental health problems. The Department for Education recommends it is used to help decision-making about links with Child and Adolescent Mental Health Services (CAMHS); suggests that ‘in the longer term, data from SDQ returns will give an indication on how effective the service provision provided is in meeting the needs of looked after children’ and notes that over time ‘records can show a child’s progress – whether difficulties identified remain or, if appropriate interventions have been put in place, whether they have eased’.

The Department for Education’s aim thus seems to be to use the SDQ in multiple ways: as an indicator of those children and young people (CYP) who are at greater risk than the general population of developing mental health problems; as an outcome measure to monitor the impact of services; and to track CYP who are in the care of the state over time. Since routinely collected demographic, health and placement variables are included with the SDQ in the data set (English SSDA903) it is also a potential source of rich longitudinal data for researchers. SDQs are not completed at entry into care, which rules out before–after analyses; however, it should be possible to use the data set to track demographic, health and/or placement
correlates of changes in scores over time. As far as we are aware, this is the first time analyses such as these have been done.

Aims

This article examines:

(a) the degree to which the current programme has achieved the intention of providing screening to identify the scale of the problem and whether it has had impact on the mental health of looked-after children in England;
(b) the potential value of analysing the data set created by that programme.

Method

Examining population trends

A common first step in evaluating screening programmes is to examine population trends (e.g. time trends of breast cancer mortality to assess the impact of mammographic screening11). The SDQ is an indicator of the prevalence of disorders.9 Therefore, one way to evaluate screening of looked-after CYP is to examine aggregate, population-level trends in the SDQ scores over time. In this case, the screening ‘intervened with’ the outcome measure and, if screening had a positive impact on practice (e.g. leading to effective referral and treatment), we might expect this to be reflected in reduced population SDQ scores over time. Publicly available aggregated data (for example13) allowed us to examine population trends in the annual SDQ returns for the 9 years (2009–2017) for which data were available. These include the number of valid SDQ returns (overall and for individual local authorities); percent of those eligible with a return; the mean total difficulties score (range 0–40); and percentages with ‘normal’ (0–13), ‘borderline’ (14–16) and ‘abnormal’ scores (17–40).

Examining representativeness

To accurately identify the scale of the problem, the SDQ data set would either need to have 100% coverage (the aim of the SSDA903 data collection) or cover a representative sample. To investigate representativeness, we conducted analyses based on the English SSDA903 data set provided, on request, by the Department for Education, which included SDQ data from 2009 to 2012. This comprised individualised data (including demographic and placement-related variables as well as the SDQ) collected annually from every English local authority relating to all CYP who had been looked after continuously for a year or longer at 31 March of the year in question. We compared selected key characteristics of children aged 4–17 with and without an SDQ to determine the representativeness of those with an SDQ data return.

Examining attrition

For meaningful longitudinal analyses (based on a data set linking individual children over 2 or more years), a representative sample needs to retain sufficient numbers over time and the characteristics of those retained compared with those lost to follow-up should be known. We selected children with a 2009 SDQ return and examined the proportion retained and whether those retained over time differed from those lost. We did this by comparing the 2009 (baseline) characteristics of children with and without longitudinal data over 2, 3 and 4 consecutive years.

Ethical approval

The study was approved by the University of Glasgow Medical Faculty Ethics Committee (2011/FM06009). Data were provided by the Department for Education Data Services Group in 2011 and 2014.

Results

Population trends

Figure 1 (based on Table 1) shows SDQ completion rates and scores from 2009 to 2017. Since the introduction of compulsory data collection, the mean SDQ score has remained consistently close to 14, with around half all children screened falling within the abnormal or borderline score categories. Levels of missing data were around 30% each year from 2009 to 2015 and 25% in 2016–2017. Table 1 shows the range of local authority data-return rates from 2009 to 2016 (2017 data by local authority not available). In 6 of the 8 years, a small number of local authorities submitted no returns; however, over this period the percentage of local authorities submitting returns for 66% or fewer eligible children decreased from 34.4% in 2009 to 21% in 2016.

Representativeness

Table 2 compares the characteristics of the CYP about whom data returns were and were not made in 2009 (results similar for subsequent years). It shows that those with an SDQ were significantly (all P < 0.000) more likely to be white (59% compared with 55%), in the middle of the age range (67% of 11- to 15-year-olds compared with only 39% of 16- to 17-year-olds), have no disability (59% compared with 45%) and to be fostered (64% compared with around 50% living in adoption, temporary or residential accommodation; 39% with parents; 17% living independently).

Attrition

Table 3 compares the 2009 (baseline) characteristics of children with and without longitudinal data over 2, 3 and 4 consecutive years. Of those with an SDQ return in 2009, 64% were retained for 2 years (2009–2010), whereas only 28% were retained for 4 years (2009–2012). Those retained in the longitudinal data set were similar to those lost in respect to gender and ethnicity, but they were significantly (all P < 0.000) less likely to have had a disability, and were more likely to have been in foster care and to have had an abnormal score in 2009. For example, those retained from 2009 to 2012 included 28% of those with no disability compared with 17% with a disability, and 30% whose 2009 SDQ scores were abnormal compared with 26% whose scores were normal. Children with SDQ returns over consecutive years are therefore not representative of those with an SDQ in any 1 year.

Discussion

Examination of the English SSDA903 data set shows no change in levels of mental health problems in looked-after children since routine screening was introduced in 2009. We found significant levels of missing data and poorly representative cross-sectional and longitudinal samples.

Whether to screen for mental disorders in looked-after children

Given these findings, a first reaction might be to ask whether SDQ screening of looked-after children is justified. Screening programmes are ‘designed to detect early signs of disease in the population and then to provide a reliable method of referral for diagnostic testing and further treatment’.15 The following ten
influential principles \textsuperscript{16} (first published in 1968 and described as a public health classic \textsuperscript{17}) have been widely used to consider whether to screen populations for noninfectious diseases: the condition should be an important health problem, there should be an accepted treatment, facilities for diagnosis and treatment should be available, there should be a recognisable latent/early symptomatic stage, there should be a suitable test/examination, the test should be acceptable, the natural history of the condition should be adequately understood, there should be an agreed policy on whom to treat, the economic costs of case finding and of providing care should be considered and case-finding should be a continuing process. \textsuperscript{18} We suggest these criteria are largely fulfilled by using the SDQ to screen for mental health problems in looked-after children. In particular, prevalence studies show high mental disorder rates within the looked-after population, \textsuperscript{10} indicating public health importance. Understandings of the natural history of child and adolescent mental disorders are increasing, with evidence that early symptoms can often be identified. \textsuperscript{19} Cost-effective, evidence-based programmes for particular groups, such as Multidimensional Treatment Foster Care \textsuperscript{20} and Attachment and Biobehavioral Catch-up for vulnerable infants, \textsuperscript{21} are available. In addition, the SDQ is a cheaper, shorter alternative to longer measures yet it has good sensitivity and specificity. \textsuperscript{22}

More recently, it has been suggested that screening programmes should be evaluated in terms of the balance between their benefits (probability of an adverse health outcome without screening; degree to which screening identifies all those who suffer the adverse health outcome; health benefit of earlier versus later treatment) and harms (frequency and experience of those with false-positive tests or who are over-diagnosed; frequency and severity

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{Summary of Department for Education SDQ aggregated data 2009–2017. Percentage with ‘normal’, ‘borderline’ and ‘abnormal’ scores (left-hand axis); mean total difficulties score (right-hand axis); x-axis shows percent SDQ returns from those eligible in each year.}
\end{figure}

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline
\hline
\hline
Number of valid SDQ returns & 22,700 & 22,810 & 23,870 & 23,480 & 23,650 & 26,020 & 27,610 & 28,810 & \\
\hline
Percentage of those eligible with SDQ returns & 68% & 68% & 69% & 71% & 71% & 68% & 72% & 75% & 76% \\
\hline
Percentage of local authorities submitting returns for: & & & & & & & & & \\
0% eligible children & 3.2% & 0.0% & 0.0% & 3.3% & 3.3% & 2.0% & 2.0% & 2.0% & \\
1–33% eligible children & 6.5% & 3.2% & 5.2% & 2.6% & 2.6% & 2.0% & 7.9% & 4.6% & 2.6% \\
34–66% eligible children & 24.7% & 27.9% & 29.2% & 23.7% & 28.3% & 25.7% & 16.4% & 16.4% & \\
67–99% eligible children & 50.6% & 61.7% & 59.1% & 65.7% & 59.9% & 56.6% & 70.4% & 75.7% & \\
100% eligible children & 2.6% & 2.6% & 1.9% & 2.6% & 5.3% & 5.9% & 4.6% & 4.6% & 2.6% \\
	extsuperscript{a} >100% eligible children & 9.1% & 0.0% & 0.0% & 0.0% & 0.0% & 0.0% & 0.0% & 0.0% & \\
Described as ‘not applicable’ & 3.2% & 4.5% & 4.5% & 2.0% & 1.3% & 2.0% & 2.0% & 0.7% & \\
\hline
Mean SDQ difficulties score & 13.9 & 14.2 & 13.9 & 13.9 & 14.0 & 13.9 & 13.9 & 14.0 & 14.1 \\
\hline
\hline
\% ‘Normal’/‘Borderline’/‘Abnormal’ & 50% & 49% & 51% & 51% & 50% & 50% & 50% & 49% & 49% \\
\% ‘borderline’ score & 12% & 13% & 13% & 13% & 12% & 13% & 13% & 13% & 13% \\
\% ‘abnormal’ score & 38% & 38% & 37% & 37% & 38% & 37% & 37% & 38% & 38% \\
\hline
\end{tabular}
\caption{Summary of Department for Education SDQ aggregated data over 7 years (2009–2017)}
\end{table}

SDQ, Strengths and Difficulties Questionnaire.

\textsuperscript{a} 2009–11 sample stated as aged 4–16.
\textsuperscript{b} 2012–17 sample stated as aged 5–16.
\textsuperscript{c} Based on a total of 154 local authorities in 2009–2011 and 152 local authorities in 2012–2017 (from 2012 Cheshire and Bedfordshire ceased to exist as separate authorities).
\textsuperscript{d} 2017 data by local authority not available.
\textsuperscript{e} In 2009 there was an anomaly with the data returns and a small number of local authorities returned more than 100% of data.
\textsuperscript{f} SDQ range 0–40; categorised as 0–13 = normal, 14–16 = borderline and 17–40 = abnormal.

\textsuperscript{488} https://doi.org/10.1192/bjo.2018.70 Published online by Cambridge University Press
of harms of treatment). Weighing up this balance in the context of screening looked-after children requires acknowledgement of the potential stigma of a mental disorder diagnosis/label; in light of this, evidence-based interventions (e.g. enhanced foster care, enhanced sensitivity to foster infants, additional resources) may seem less likely to cause harm than treatments for screening-identified physical illnesses (e.g. surgery, radio- and chemotherapy for cancers). Again, screening for mental health problems in looked-after children appears justified.

Do we have an effective screening programme in England?

Identification of those with problems is only the first step; the next is to address those problems. However, English local authorities have inadequate referral routes to CAMHS once the SDQ has identified children with possible mental disorders. The current programme of compulsory SDQ returns comes at a time when financial pressures mean, many specialist teams offering support to looked-after children falls short, due to large amounts of missing data and no link to any Department of Health and Social Care accept that looked-after children should undertake the SDQ as a starting point when they come into care, and then each year as part of compiling an accurate picture of their health needs’ (p. 6).

What could the data tell us beyond screening? This mass data collection exercise might be useful for examining geographical variations or time trends, or as a performance indicator for local authorities. Data derived from the SDQ screening programme are only available for about 70–75% of children in any 1 year and approximately 40% of children move in and out of the care system each year; meaning that useful, representative, longitudinal analyses would be challenging – although not impossible if levels of missing data were reduced. This vast and annually increasing data set has great research potential; it is possible that lack of change at the population level masks real effects at an individual level and careful consideration of how individual analyses could be achieved should be part of any revision of the system.

Suggestions for improvement of the current system We suggest consideration of the further opportunities the annual SDQ data collection affords, both in terms of its analytic potential and as a screening programme. Currently the screening programme falls short, due to large amounts of missing data and no link to any
‘next steps’ for those children whose scores indicate likely disorder. As a data set, investment in better completion and more complex analyses may increase understandings of (likely reciprocal) associations between looked-after children’s emotional/behavioural difficulties and both demographic and placement-related factors. Screening should not occur in isolation; investment in better systems would ensure SDQ scores for individual children are scrutinised, used in decision-making and – where they indicate likely psychiatric diagnosis – trigger clear referral pathways. These actions could result in improved placement and health outcomes for looked-after children, and this would be a worthwhile investment.

Limitations

Our use of publicly available data and simple statistical analyses aimed to demonstrate time trends and examine representativeness. Some might argue that it is impossible to evaluate the impact of screening using the SDQ without conducting longitudinal analyses (e.g. comparing outcomes for those with/without an SDQ, or those coming into the system at earlier/later time points) or by examining proxy data on service referral rates, access and/or effectiveness as outcome indicators. We contend that examining population-level trends in SDQ scores offers insight into the impact of screening looked-after children, and that there are flaws inherent in any longitudinal analyses of incomplete data. Our simple analyses are thus an important first step in examining the SDQ screening programme.

Table 3  Characteristics of those with and without SDQ data for 2, 3 and 4 consecutive years

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Overall</td>
<td>22681</td>
<td>14515</td>
<td>8166</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>12857</td>
<td>66.3%</td>
<td>33.7%</td>
</tr>
<tr>
<td>Female</td>
<td>9824</td>
<td>63.5%</td>
<td>36.5%</td>
</tr>
<tr>
<td>χ² (significance)</td>
<td>1.6 (0.209)</td>
<td>4.4 (0.036)</td>
<td>1.9 (0.168)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>18059</td>
<td>64.5%</td>
<td>35.5%</td>
</tr>
<tr>
<td>Black and minority ethnic</td>
<td>4622</td>
<td>62.0%</td>
<td>38.0%</td>
</tr>
<tr>
<td>χ² (significance)</td>
<td>9.7 (0.002)</td>
<td>11.6 (0.001)</td>
<td>1.6 (0.201)</td>
</tr>
<tr>
<td>Age in 2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>589</td>
<td>54.2%</td>
<td>45.8%</td>
</tr>
<tr>
<td>5–10</td>
<td>6344</td>
<td>70.6%</td>
<td>29.4%</td>
</tr>
<tr>
<td>11–15</td>
<td>11623</td>
<td>76.4%</td>
<td>23.6%</td>
</tr>
<tr>
<td>16–17</td>
<td>4125</td>
<td>20.2%</td>
<td>79.8%</td>
</tr>
<tr>
<td>χ² (significance)</td>
<td>4363.6 (0.000)</td>
<td>3811.4 (0.000)</td>
<td>2300.6 (0.000)</td>
</tr>
<tr>
<td>Disability in 2009a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>21908</td>
<td>64.4%</td>
<td>35.6%</td>
</tr>
<tr>
<td>Any</td>
<td>773</td>
<td>53.3%</td>
<td>46.7%</td>
</tr>
<tr>
<td>χ² (significance)</td>
<td>39.7 (0.000)</td>
<td>48.6 (0.000)</td>
<td>46.4 (0.000)</td>
</tr>
<tr>
<td>Placement in 2009b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adoption and temporary placement</td>
<td>603</td>
<td>18.4%</td>
<td>81.6%</td>
</tr>
<tr>
<td>Foster</td>
<td>17840</td>
<td>69.0%</td>
<td>31.0%</td>
</tr>
<tr>
<td>Residential</td>
<td>2888</td>
<td>52.9%</td>
<td>47.1%</td>
</tr>
<tr>
<td>Parents</td>
<td>1106</td>
<td>49.0%</td>
<td>51.0%</td>
</tr>
<tr>
<td>Independent</td>
<td>228</td>
<td>7.0%</td>
<td>93.0%</td>
</tr>
<tr>
<td>χ² (significance)</td>
<td>1321.8 (0.000)</td>
<td>1019.5 (0.000)</td>
<td>796.7 (0.000)</td>
</tr>
<tr>
<td>SDQ score in 2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal range</td>
<td>11344</td>
<td>61.7%</td>
<td>38.3%</td>
</tr>
<tr>
<td>Borderline range</td>
<td>2793</td>
<td>63.8%</td>
<td>36.2%</td>
</tr>
<tr>
<td>Abnormal range</td>
<td>8544</td>
<td>67.1%</td>
<td>32.9%</td>
</tr>
<tr>
<td>χ² (significance)</td>
<td>62.0 (0.000)</td>
<td>40.2 (0.000)</td>
<td>43.9 (0.000)</td>
</tr>
<tr>
<td>SDQ difficulties score in 2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean score</td>
<td>13.9</td>
<td>14.3</td>
<td>13.2</td>
</tr>
<tr>
<td>F (significance)</td>
<td>87.0 (0.000)</td>
<td>63.9 (0.000)</td>
<td>61.8 (0.000)</td>
</tr>
</tbody>
</table>

SDQ, Strengths and Difficulties Questionnaire.

a. Disability was defined as the reason for entry into care rather than whether or not the child has a disability. It is therefore likely to only identify children who have profound needs.

b. There were 16 cases with missing data on ‘placement’, these were excluded.

References


2. Information for researchers and professionals about the Strengths & Difficulties Questionnaire. Youth in Mind (http://www.sdqinfo.com).

