Title: Linking health and education data to plan and evaluate services for children.

Johnny Downs\textsuperscript{1,2}, Ruth Gilbert\textsuperscript{3,4}, Richard D. Hayes\textsuperscript{1,2}, Matthew Hotopf\textsuperscript{1,2} and Tamsin Ford\textsuperscript{5}

\textsuperscript{1} Psychological Medicine Department, Institute of Psychiatry, Psychology, and Neuroscience, King’s College London, London, United Kingdom
\textsuperscript{2} NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust, King’s College London, London
\textsuperscript{3} Farr Institute of Health Informatics Research London, London, UK
\textsuperscript{4} Children’s Policy Research Unit, UCL Institute of Child Health, London, UK
\textsuperscript{5} Child Mental Health Research Group, University of Exeter Medical School, UK.

* Correspondence to: Dr Johnny Downs, Box 63, Department of Psychological Medicine, Institute of Psychiatry, King’s College London, De Crespigny Park, SE5 8AF London, UK. Phone: +44 (0)20 3228 8553 Fax: +44 (0)20 3228 8551, Email: johnny.downs@kcl.ac.uk.

Summary: In this leading article, we report on the first area-based linkage in England between mental health, hospital and school data, covering a total population of 1.25 million. We describe this resource, give examples of how it is being used to improve public services for children, and discuss what is needed to implement this approach more widely across the UK.
**Introduction.**

Linkage of routinely collected data from public services has the potential to improve how local health, education, and social care are delivered to children. All mental health services, hospital-based child health services, schools and child protection services which serve the same local area, could be more efficient if the design, monitoring, targeting and integration of services were based on data. Health services need evidence from the populations that they serve to plan care and know whether they are meeting children’s needs, duplicating effort, or allowing some children to fall through the net. In this paper we describe how the CRIS programme has joined up data from health, education and social services for children living in four local authorities in South London to create two datasets. One linking hospital to children’s mental health services and the second linking mental health data to education data. We describe these resources, give examples of how they are being used to improve services, and discuss what is needed to implement this approach more widely across the UK.

**What data are available?**

Across England, all NHS health and state education services for children routinely generate administrative data, but few areas have managed to join these data systematically to evaluate how services could better serve their populations. Details of every NHS hospital inpatient admission, emergency department and outpatient contact are centrally collated by NHS Digital.¹ Demographic and socio-economic data on every child in state education are submitted by all state maintained schools to the Department of Education, along with information on school attendance, attainment, exclusion, child protection involvement, and special needs.² Centrally collected child mental health data has yet to become available, but nearly all local services collect this data within their electronic health record systems.³ A big challenge is meeting the technical and governance requirements that safeguards sensitive child data, but also permits the linkage across public service data resources. This challenge has been addressed by the NIHR biomedical research centre at the Maudsley and there is the potential to extend our approach to other sites.
Ten years ago, the Maudsley NIHR biomedical research centre set up the Clinical Record Interactive Search (CRIS). CRIS is a secure data repository for NHS mental health services covering four local authorities. CRIS developed a process to anonymise electronic health records for service evaluation and clinical research. The system operates strict governance controls with external service-user led oversight.\textsuperscript{4,5} CRIS retains patient identifiers, such as NHS number, separately from patients’ mental health records, which are anonymized.

CRIS has linked mental health data to education and hospital data. It took 3 years, from first application, to obtain permissions to do this from the Health Research Authority, NHS Digital and the Department for Education. The linkage process itself involved CRIS sending patient identifiers (names, and dates of birth, postcodes), without any mental health information to NHS Digital and to the Department for Education, where the identifiers were linked, and data from education and hospitals were de-identified and returned to CRIS. The CRIS secure environment now holds two linked datasets, education data linked to mental health data, and a second dataset containing mental health and hospital data. These datasets are kept separately, with all identifiers (names and NHS or pupil ID numbers) removed.

The CRIS system covers all NHS mental health services for four local authorities, which service a population of 1.25 million people. Patients using mental health services are made aware of how their data are used through notices in clinics, websites and regular public engagement events. Although patients are not asked for consent to use their data for service evaluation or research, they are able to opt out. Only three individuals have asked to opt out of CRIS in six years. In 2014, the CRIS system was extended to four more mental health trusts (in 16 local authorities) and could be extended beyond mental health to other services.\textsuperscript{6}
Using linked data from schools and mental health services

The population

The linked schools and mental health dataset captures data for approximately 160,000-190,000 children each year from 2007-13. To be included, children need to be aged between 4 and 16 years (see figure 1 for population numbers) and be resident in Southwark, Lambeth, Lewisham or Croydon. These areas are culturally and economically diverse, representing both outer and inner London regions. The catchment population has substantially higher proportions of families from black minority ethnic groups and/or born outside UK compared with rest of London and England. Highest and lowest socioeconomic groups are overly represented compared with England; with higher rates of unemployment, but also higher levels of education. Linkage with the national pupil dataset means that information on education is still captured for those attending state school outside the local catchment area, and for those who move in or out of the area. Some of the population are not routinely captured. Children attending independent (meaning private) primary schools are not represented (~5% of the population aged under 12). Children attending independent secondary schools are included when they sit any national examinations (e.g. GCSE or A level).

What can be measured?

Alongside socio-demographic characteristics, the national pupil dataset provides rich information on childhood development. It tracks indicators of cognitive ability via routine teacher based assessment of language and numerical ability as children start school, and then via standardized academic assessments in mid and late childhood. It captures indicators of special educational needs such as physical problems, including deafness and visual impairment; emotional and behavioural problems, and autism spectrum disorder and learning disability. The dataset also captures episodes of children being excluded from school and indicators of absenteeism. Children’s social care data has also been linked, which includes social service referrals and investigations, including details of children who are
placed into out of home care.

The CRIS system enables researchers to access electronic mental health record data for approved studies.\(^5\)\(^9\) Data available for research includes structured information (e.g. data entered by clinicians from drop down lists) such as past and present ICD-10 psychiatric diagnoses, appointments attended, and routine outcome measures (e.g. Strength and Difficulties Questionnaires)\(^10\), and risk assessment details including risk of self-harm, self-injury, aggression to others.\(^11\) Natural language processing software is used to enhance this data by extracting information predominately found in clinical progress notes and correspondence that might include more detail about family mental health problems, substance misuse, pharmacotherapy, and symptoms.\(^9\)

**How can school and mental health data be used to improve services?**

The linked school and mental health data has many potential applications. It can provide detailed information on patient pathways and the extent of inequalities to services. This information can be used to flag gaps in existing healthcare provision and direct where new services are needed. National and local surveys have provided consistent evidence that timely access to services varies by social status and area of residence.\(^12\) Data suggest that young people at high risk for mental health problems, looked after children and care leavers, those at risk of social exclusion or who have experienced abuse, or with long term physical health conditions, are the ‘hardest to reach’ and more likely to receive insufficient or fragmented care.\(^13\) Because, local areas have considerable flexibility in how they commission child mental health services, there is a risk that ‘hard to reach’ groups are least likely to receive services. Mental health-school linked data can help understand which children receive support amongst socially vulnerable groups in each local area. Using data in this way to map service provision is particularly pertinent for integrated child health programmes which aim to tackle the potential inefficiencies and inequalities of current condition-specific pathways.\(^14\) At present, local
areas have very limited information on how mental health resources are accessed by vulnerable children, which include looked after children, those with a history of social services contact, prolonged absences from school, permanent exclusions, and with complex education needs. Using these linked data, we can gain a clearer picture of how well education and mental services overlap to address emotional and behavioural difficulties, and the shared awareness of special educational needs across both services.

Linkage of schools data to mental health services also offers opportunities for targeting school based prevention strategies. For example, the funnel plot in Figure 2 shows variation between mainstream schools in referrals to child and adolescent mental health services in the four local areas for children aged less than 8 years. Outliers on the funnel plots are of particular interest and warrant further exploration: very high rates could reflect high levels of population need and/or school-wide difficulties in managing emotional and behavioural problems, or conversely, very low referrals to mental health services may reflect excellent in-school support and provide a model of good practice. Using similar techniques, the data can be used to examine whether potentially more ‘contagious’ adolescent mental health problems like eating disorders, self-harm or suicidal behaviours cluster within schools. Findings can then be used to prioritise schools for preventive strategies. There is also a need for research to examine associations between educational achievement, self-harm presenting to mental health services, and the potential impact of school based interventions, as almost no research has been conducted on this topic in the UK.

[figure 2 about here]

Using linked hospital-mental health service data to inform services
There are a number of policy-relevant research areas that can benefit from using linked mental health and hospital administrative data. One example is the evaluation of policy initiatives to improve the quality of crisis care for young people. There are approximately 200,000 episodes of self-harm that present to emergency services each year in the UK, with the highest rates amongst adolescents and young adults. Between 25–50% of adolescents presenting to emergency care with self-harm do not attend any follow-up mental health support. Emergency departments have an important influence on future engagement with treatment. By adapting an approach developed in adult populations, we can track temporal shifts in rates of emergency department attendances for self-harm or suicidal behaviour for the 40,752 children and adolescents seen each year from the four local authorities served by CRIS. We can assess whether practice changes in emergency departments result in reduced rates of attendance for self-harm in the long term.

Another example is the use of linked hospital-mental health data to follow up children hospitalised with long term conditions to investigate their use of mental health services and psychiatric co-morbidity. We can evaluate the types of patients who receive mental health care, when, and which factors are associated with treatment gaps and/or reliance on emergency care and unplanned admissions. These studies can provide information on whether systems of care need to change to ensure particular populations with chronic health problems, such as ethnic minorities or socially disadvantaged children, receive equitable access to mental health services.
CRIS: a sustainable resource for evaluating child health policy and service improvement

The CRIS system offers a sustainable resource for population-based analyses of linked patient level data to inform child mental health and acute hospital services and education services. Because CRIS uses data extracted from electronic record systems it provides a powerful platform for continuous evaluation of local child health policy initiatives. The CRIS system provides an efficient, area-based resource for research, service planning and evaluation with patients followed up across the country. CRIS is being reproduced in other areas, potentially leading to a number local areas having fine-grained information to better target local resources. However, there is still considerable work to be done. Health commissioners and other decision makers at local and national levels will need to develop sustainable means of implementing the knowledge which resources such as CRIS can deliver. This is essential if we wish to complete the Learning Health System cycle, and use our informatics resources to drive healthcare improvement and innovation. Alongside this, public engagement, understanding and support is vital. If we want to adopt these systems further families, child health advocates, academics, clinicians and policy makers will need to decide together how local linked resources are best safeguarded and used in commissioning services.

We hope in time that others will be encouraged to extend the CRIS model to link data for children across public services. By doing so we hope to reduce the unmet need among vulnerable children and to move the discussions on from ‘not knowing’ to accurate and responsive information on which to base public health strategies for children and young people.
Acknowledgments

The authors thank Robert Stewart and Matthew Broadbent for their comments on previous drafts. We also wish thank NIHR Maudsley BRC CRIS Team especially Hitesh Shetty for supporting the data extraction for HES-CAMHS data. In addition, the authors are very grateful to Richard White, Karen Stevens, and Martin Johnson and members of DfE National Pupil Database team, who provided invaluable support throughout the project.

Contributors: JD, RG and TF conceived and drafted the article, with additional critical revision from RH and MH. JD wrote analysis plans and analysed the data.

Funders: This paper represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. Other funders include Guy’s and St Thomas’ Charity and the Maudsley Charity (BRC-201110035). Dr Downs is supported by an Medical Research Council (MRC) Clinical Research Training Fellowship (MR/L017105/1). Dr Hayes was funded by an MRC Population Health Scientist Fellowship (MR/J01219X/1). Professor Hotopf is a NIHR Senior Investigator. Professor Gilbert is a member of the Farr Institute for Health Informatics Research, London (MRC grant numbers: London MR/K006584/1)

Disclaimer: The views expressed are those of the authors and not those of the NHS, the NIHR, the Department of Health or Department for Education. The study sponsors played no part in the design, data analysis and interpretation of this study; the writing of the manuscript or the decision to submit the paper for publication, and the authors’ work was independent of their funders.

Competing interests: None declared.
References


Figure 1
Linked data resources to provide an anonymised multiagency dataset covering child and adolescent mental health services, hospital attendances, education services and social service activity in South London. CRIS, Clinical Record Interactive Search; NHS, National Health Service.
Figure 2
Plot showing referral rates to Child and Adolescent Mental Health Services for each school by Key Stage 1 (infant school) as a function of the pupils enrolled between 2008 and 2013. The average referral rate is 4.3% (shown as a horizontal line). Control limits are also plotted above and below this mean.