

**Drawing on Multiple Sources for Developing a
Preliminary Payment System Framework
for Child Mental Health Provision**

Vostanis Panos, Professor of Child Psychiatry, University of Leicester

Martin Peter, Module Convenor Research Methods and Data Analytics Lead, Evidence Based Practice Unit, Anna Freud Centre, London

Davies Roger, Clinical Psychologist, City and Hackney CAMHS, East London Foundation Trust

De Francesco Davide, Lecturer in Statistics, University College London

Jones Melanie, Improvement Programme Lead, Evidence Based Practice Unit, Anna Freud Centre, London

Sweeting Ruth, Clinical Psychologist, City and Hackney CAMHS, East London Foundation Trust

Ritchie Benjamin, Pilot Site Manager, Evidence Based Practice Unit, Anna Freud Centre, London

Allen Pauline, Reader in Health Services Organisation, Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine

Wolpert Miranda, Director, Evidence Based Practice Unit and CAMHS Outcomes Research Consortium, Anna Freud Centre, London

Correspondence: Professor P. Vostanis, University of Leicester, Greenwood Institute of Child Health, Westcotes House, Westcotes Drive, Leicester LE3 0QU; tel: 0116 2252885; fax: 0116 2252881; e-mail: pv11@le.ac.uk

Word count: 4,994

Keywords: provider payment, child mental health services, case-mix, episode of care

Abstract

Objectives: There is an urgent requirement to develop a payment system for services for children with mental health problems that allows more targeted purchasing based on ‘fairness’ and ‘need’. This is currently constrained by lack of clinical consensus on best ways forward, wide variation in practice, and lack of data about activity and outcomes. In the context of a national initiative in the UK, the aim was to use a mixed methods approach to develop a framework that could form the basis for an improved payment system.

Methods: Three inter-related studies addressed: a) qualitative consultation with child and adolescent mental health services (CAMHS) stakeholders on what the key principles for establishing a payment system should be, via online survey (n=180) and participatory workshops (n=91); b) review of relevance of all national clinical guidelines (n=15); and c) a quantitative case study on the relationship between disorders and resource use (n=1,774 children from 23 teams).

Results: CAMHS stakeholders stressed the need for a broader definition of need than only diagnosis, as well as the measurement of indirect service activities and the inclusion of appropriate outcome measurement. National clinical guidance suggested key aspects of best practice for care packages, but did not include consideration of contextualised factors such as complexity. Modelling case study data found that problem type and degree of impairment independently predicted resource use, alongside evidence for substantial service variation in the allocation of resources for similar problems.

Conclusions: A framework for an episode-based payment system for CAMHS should include consideration of complexity factors and indirect service activities; support evidence-

based care packages; consider different needs in terms of impairment and symptoms; and include outcome measurement as a core component.

Background

Case-mix adjusted episode-based payment

There is considerable international interest in the development of ‘episode of care’ units for use in payment systems for secondary health services. A unit of activity of this sort is a series of temporally contiguous services delivered by a single or multiple providers, to a particular type of service user, as defined by the ‘case-mix’ group the service user is categorised into.^{1, 2} For the purposes of this paper we use the term ‘episode’ to include both periods between admission and discharge from a service, and specified periods of time (longer than a day) in receipt of ongoing care.^{2, 3} Application of episode of care units to reimbursement has spread worldwide since the introduction in the 1980s of Medicare’s Prospective Payment System (PPS), which uses diagnosis-related groups (DRGs) to account for case-mix differences in hospital in-patient episodes. Structuring payment in this manner is appealing to policy makers and purchasers in tax or insurance funded health systems for several reasons discussed below. Research has concentrated on the PPS and its international variants, as similar payment systems in mental health are currently under early development.⁴

Quantity and cost of health care services

With regard to encouraging the supply of an appropriate quantity of health care, it is helpful to consider the rationale for introducing episode-based payment in the context of existing funding arrangements. If providers’ costs are reimbursed according to individual items or components (e.g. per diagnostic procedure, per day of in-patient care, per out-patient appointment), aggregating these into episodes with prospectively set prices is predicted to improve productivity and control costs at episode level.⁵ The argument is that a prospective price assigned to an episode of care creates a financial incentive to reduce the quantity of

items or components of care within it, thereby countering the tendency for ‘excessive’ components in the presence of health insurance. There is also an incentive to substitute lower-cost for higher-cost components, where feasible.⁶

In systems historically funded through block budgets at the purchaser-provider level, moving to episode-based activity payment creates closer linkage between demanded and supplied quantity of health care. This may help to reduce waiting times by increasing throughput, while also providing a safeguard against the potential problem of over-provision of items within episodes.^{7, 8} However, here lies a challenge with episode-based payment – although there are incentives to control quantity and costs within episodes, there is also an incentive to increase the number of episodes.^{2, 9} This may create undesirable consequences for the health system.⁷

The effect of introducing episode-based payment with prospectively set prices on hospital productivity has been explored in several countries by examining average length of stay (LOS), a proxy for the quantity of inputs or costs required to produce an episode. Studies have generally found the introduction of this reimbursement approach to be associated with modest or no reductions in average LOS.^{5, 8} Caution in interpreting results is warranted due to the challenge of isolating the effects of payment system reform from concomitant policies and trends in hospital productivity.⁸

Quality of care

The contribution of episode-based units of activity to quality improvement in terms of the safety or effectiveness of care is envisioned through several mechanisms.^{2, 10-12} For example, categorising episodes according to their complexity or severity identifies ‘products’ or ‘outputs’ of a provider that bear greater relation to users’ health problem(s) than traditional conceptualisations of outputs such as the number of (unadjusted) discharges or out-patient attendances.¹⁰ This permits inter-provider comparison of the average costs/LOS of particular episode-based ‘products’, or identification of episodes with unexpectedly high or low costs/LOS within a given product category, which may indicate where quality deficiencies lie.^{10, 11} This kind of analysis can be undertaken independently of a payment system, although an overarching financial agenda may encourage it, and may stimulate improvements in the accuracy and completeness of the coded data it relies on.^{6, 11} However, some commentators argue that case-mix classifications designed to predict episode costs for reimbursement purposes have only limited use for quality assurance, either due to high residual diversity of patients within particular categories,¹³ or their inability to provide satisfactory information on predicted health outcomes.¹⁴

Another potential path to quality improvement applies to countries where existing funding arrangements were such that providers’ income was independent of the numbers of patients treated (block contracts). By allowing patients to choose their provider and linking providers’ revenue to the number of episodes delivered, it has been proposed that providers would be encouraged to compete for patients on the basis of quality.^{12, 15} This relies on patients receiving and using accurate and interpretable information on the quality of care of specific providers.⁹

Conversely, it has been widely expressed that reimbursement structured by episode-based units of activity could have a negative impact on quality, due to the aforementioned incentive to limit the costs of episodes.^{5,9} This raises the possibility of providers reducing necessary services ('skimping'), or discharging patients too early, particularly if the severity or complexity of certain conditions is inadequately accounted for by the payment system.¹⁶ Empirical study of effects on quality has been limited by deficiencies in the availability of outcome data, hence no firm conclusions can be drawn.⁵ The limited findings do, however, broadly suggest that in the acute hospital sector, such payment systems have not led to a dramatic - negative or positive - effect on outcomes.¹⁶

Fairness and 'need'

In payment system design, classifying episodes by case-mix aims to produce groups of episodes with similar costs, on average, for the same quality of care.^{1, 10, 13} It follows that if prices assigned to episode categories accurately reflect their relative average costs, then, all else being equal, providers treating a more resource intensive population should receive a higher level of reimbursement. Fairness in payment under this system, therefore, rests in part on how well the grouping methodology accounts for variation in the costs of episodes attributable to patient characteristics that are unevenly distributed among providers.^{13, 17}

In the context of mental health payment system development, fairness has also been considered from the perspective of the 'needs' of service users.^{4, 18-20} It is hoped that by categorising them users according to a particular logic applied to their attributes at the start of episodes (as opposed to the interventions they received during episodes), the resultant

episode classifications will provide a forward-looking estimate of their relative health care needs. A reimbursement system with episode prices weighted according to need could, therefore, have positive implications for equity of access, especially if the preceding payment arrangements take minimal or no account of need. However, it is useful to note the concerns arising from hospital case-based payment that a ‘residual’ range of complexity or severity of cases within a payment category may lead providers to develop strategies to avoid treating the more complex and severe cases within that category, for fear that payment will be inadequate.^{6,9} If these ‘above average’ cases cannot be identified early enough to be avoided, when they become apparent through higher than average resource use, the quantity of their care might be reduced. In designing needs-based episode categories for mental health, it is conceivable that the more precise the weighting of the episode categories for ‘actual’ need, the lower the risk of decreased access to care for the ‘higher need’ service users within episode categories. Rewards or sanctions linked to the quality of care and additional payment for high cost outliers may also assist with mitigating this risk.^{6,9}

Consequently, a major challenge in designing classifications of this type is the quantification of need. ‘Need’ is a debated concept, which arguably can vary depending on different perspectives and criteria for its nature and extent.²¹ It is, therefore, unsurprising that there have been significant differences in the approaches to designing such classifications, with developers drawing to varying degrees on clinical judgement (‘normative need’) and statistical analysis of the relationship between service user variables and resource use (as a proxy for need).^{18,19,22} In regards to applications of methodology incorporating analysis of service user–resource use relationships, reported constraints have been the provider-level variation in treatment activity, and how little the information collected on patient characteristics at the start of an episode can predict variance of treatment activity – the ‘weak

signal-loud noise' problem.^{3, 18} For adult mental health services in England, a combined data analysis and clinical judgement approach for determining case-mix categories was eventually settled on, specifically employing a statistical method of cluster analysis (not including resource use data) and multidisciplinary panel review of data.²²

Developing a payment system for CAMHS in England

This paper reports on the early stages of a project aimed at creating a set of episode of care units to underpin a payment system for child and adolescent mental health services (CAMHS). It arose from a policy commitment to extend the move away from block contracts into more controlled and transparent 'units of purchase' beyond acute physical health services.²³ Currently out-patient CAMHS are purchased mainly by block contracts, or contracts based on itemised units of activity such as face-to-face contacts.²⁴ In part the wish to move from block contracts was fuelled by concerns already outlined above such as that these arrangements do not appropriately take into account numbers of service users cared for or their needs.²⁵ In addition, it was hoped that since the move from block contracts would involve agreeing a national way of categorising mental health care into 'units of purchase' that are clinically similar and have similar resource needs, this would also facilitate greater understanding of the quantity and quality of care delivered, since it was recognised that currently data are largely missing in regards to these crucial dimensions²⁰. Moreover, it was hoped that the new payment systems would support more productive discussions between payers and providers, and comparison of providers' performance ('benchmarking').^{4, 20}

An initial step in the process of designing such a system is to develop a framework for classification with the desired property of differentiating groups of service users with different levels of health care need. The focus of this paper is the process undertaken to establish and start applying the classification development methodology. Aspects of CAMHS likely to pose particular difficulties are limited evidence on effective interventions in real settings, mainly for children with complex problems;²⁶ limited validation of assessment and outcome measures;²⁷ and lack of data on ‘indirect’ activities (related to a service user’s care but where s/he is absent). In light of these and other challenges with quantifying need, the aim was to consider and integrate both normative and empirical knowledge, by drawing on different sources.

Methods and Results

This preliminary phase consisted of three inter-related studies:

1. Consultation with CAMHS stakeholders.
2. Review of clinical guidelines on effectiveness of interventions.
3. Case study on association between clinical problems and resource allocation.

Study 1: Consultation with CAMHS stakeholders

The principles that should drive the development of a payment system for CAMHS were explored with a sample of stakeholders through a nationally promoted online questionnaire survey and two participatory workshops. The aim was to establish views on: a) should payment case-mix groups be diagnostic, needs or resource led; i.e. should groups be defined by common diagnostic characteristics, broader needs, or the level of resources used irrespective of the former; b) whether and how complexity factors should be taken into consideration; and c) should non-face-to-face activities be included in the model? The online

questionnaire survey was promoted through the project website and CAMHS networks in England. The 180 participants consisted of clinicians (66%); service managers and finance officers (19%); voluntary sector (7%); commissioners (6%); social care providers (2%); and service users (2%). These worked in a wide spread of English regions. The questionnaire included open questions, which were subjected to content analysis by two raters (MJ and PV). The findings were presented at two participatory workshops. Attendants were invited through the same procedure. Of the 91 participants (57 and 34 at each event), approximately two-thirds (64.8%) were again clinicians. The group facilitators collected individual and group recommendations, which were subjected through the same process of content analysis.

The key emerging categories were, that payment systems should adopt the following principles: needs are broader than diagnoses, taking complexity factors into consideration; needs should be linked with resource use and outcomes; indirect activities should be taken into account and costed; outcome measures to support clinical practice but not determine payment as such; allowance of migration across case-mix groups over time, including mapping to adult groups at 18th birthday; validation of non-burdensome assessment tools; and anticipation of varying degrees of joint health and social care commissioning.

Study 2: Contextual review of clinical guidelines

Clinical guidelines of the National Institute of Health and Care Excellence (NICE) were reviewed, with the aim of identifying links between evidence-based care packages and contextual factors, i.e. whether interventions should be adapted in terms of resource allocation to address complex needs. NICE guidance is based on systematically reviewed literature collated into a report.²⁸ Eleven guidelines were specific to children (depression,

PTSD, obsessive-compulsive, self-harm, conduct, ADHD, autism, eating disorders, alcohol dependence, looked after children, and well-being in education), and four guidelines for adults referred to children (anxiety, bipolar, psychosis, borderline personality disorders). These were reviewed in terms of key treatment recommendations, how/where interventions should be implemented, length, skill mix, costs, contextual factors (risk, engagement, complexity), and expected outcomes. Complexity was defined as a multitude of mental health and social care factors involved, i.e. this was distinct from symptom severity. Guidelines were independently reviewed according to these criteria on a rating template, initially by one reviewer (RS), and were subsequently revisited by two reviewers (PV and RD).

All guidelines specified recommended treatment modalities and their duration, based on the available treatment in the area. There was variation in the definition of practitioners' skills and care levels (referred to by eight guidelines). Contextual factors were usually mentioned in terms of comorbidity (e.g. ADHD), with some reference to environmental risks. Six guidelines defined desired outcomes in terms of symptoms and impairment, but not specifically to care packages. Seven guidelines referred to costing, with three (depression, conduct, ADHD) providing details on care packages, including indirect costs. Overall, identified conditions were predominantly diagnostic-driven, but usually not contextualised (framed in relation to real practice and settings) in planning appropriate interventions. The limited evidence from trials in clinical settings necessitated that findings were complemented by expert consensus, i.e. RCT findings only constituted a moderate proportion of the evidence. Indirect service activities such as consultation and liaison were referred to, but not specified.

Study 3: Case study on association between disorders and resource allocation

A desirable feature of a case-mix classification is its ability to identify groups that are as internally homogeneous as possible with respect to resource use, remaining mindful that resources utilisation does not necessarily reflect resource need. We analysed CAMHS clinical records to establish the extent to which resource use could be predicted from information on problem type and severity. Anonymised clinical records of closed cases were provided by an English inner-city CAMHS provider. All children were seen in the community, although a small number may have also had episodes of in-patient treatment. Interventions were delivered by 23 CAMHS teams across three localities. The sample (n=1,774) was constructed by selecting children for whom valid information was available on gender, age, presenting problem, impairment (measured by the Children's Global Assessment Scale - CGAS²⁹) and number of appointments attended. The sample sizes in the three localities were 500, 590 and 684, respectively. The number of children in the 23 teams ranged from 1-350 (median=38). Of those, 961 (54%) were male and 813 (46%) female. The age distribution was: under 5 years 144 (8 %), 6-12 years 789 (44 %), and 13-18 years 841 (47 %).

Presenting problems were recorded by a clinician at assessment in accordance with data specification of the Child Outcomes Research Consortium (CORC).³⁰ It was possible to record the same child as having more than one problem. We initially considered 22 different problem types, which occurred in at least 0.4 % of the sample, plus one 'residual' category for less frequent types. For the purpose of the presentation, problem types were categorised into four groups, based on significance tests of coefficients in the statistical model and similarity of problem type: resource use (psychosis, eating disorder) 42 (2%); moderate resource use related to emotional problems (emotional problems only or comorbid with conduct, eating, self-harm, or learning disability) 764 (43%); moderate resource use related to

conduct problems (conduct problems only, or comorbid with other than emotional problems) 104 (6%); and low resource use (all other types of problems and problem combinations).

The dependent variable was the number of appointments attended after assessment. This is a discrete variable, with large variance relative to the mean. Its distribution approximated the negative binomial probability density function. Data were analysed using mixed negative binomial regression. The effects of gender, age and CGAS rating (higher scores indicate better psychosocial functioning) on number of appointments were modelled as fixed effects. To take account of the nested structure of the data (children were treated within clinical teams, which were located within one of three localities), a random intercept term was fitted for the three localities, as well as for the 23 teams. We suggest the three localities can stand as examples for CAMHS. Generic clinical teams were similar across localities, but specialist teams (e.g. for neurodevelopmental disorders) varied within each locality. The average number of appointments differed considerably between localities.

Table 1 shows the model coefficients. The interaction effect between problem type and CGAS rating was tested, but did not significantly improve model fit and was, therefore, not included in the final model. Raw coefficients do not have an intuitive interpretation, but we can interpret their signs and relative sizes. The positive coefficients show that, controlling for all other variables in the model, girls, older children and children in the ‘high resource use’ group (psychosis and eating disorder) attend most appointments.

Insert **Table 1** here

The negative coefficient of CGAS rating shows that the higher the clinician-assessed functioning at assessment, the lower is the expected number of appointments (again, controlling for all other variables). Thus CGAS rating and problem type each make an independent contribution to the prediction of number of appointments. Figure 1 shows the number of appointments by CGAS rating, as well as the predicted number of appointments from the regression model, for teenage girls (plots for girls of other ages or for boys would look similar, except that predicted values would be slightly lower). It illustrates how poor the model fits overall, since the individual values are scattered widely around the lines of best fit.

Insert **Figure 1** here

Finally, we tested whether random effects were necessary in the model. Removing the random effects led to significantly worse fit to the data (likelihood ratio test statistic $\Lambda=400.9$, $df=2$, $p<0.001$). This means differences between teams and localities cannot be explained by case-mix differences; at least not insofar as such differences are reflected in the children's gender, age, presenting problems and CGAS ratings. Clinical assessment can provide some information that is relevant for the prediction of resource utilisation. However, a large amount of variation in resource use is not predictable using the data we had at our disposal.

Discussion

In this preliminary work, we aimed to establish a framework of key evidenced principles to guide the development of an episode of care classification to support a child mental health services payment system. An ideal property of the classification is to differentiate groups of service users on the basis of their relative needs for health care. In the absence of a single

reliable approach, we utilised three sources of information. Each contributed to determining the principles, but was also faced with limitations (Table 2). The overarching conclusions indicate the importance of defining children's mental health needs in broader than diagnostic terms, including consideration of complexity; relating needs to resource use and outcomes, rather than merely activities; and considering costs of indirect, as well as direct, staff activities.

Despite the diagnostic focus of the clinical guidelines, their coverage of the majority of child mental health problems and recommendation of care packages, including suggested numbers of appointments, will make a useful contribution to estimating relative needs, and to understanding gaps between 'best practice' and current resource use. The latter point is important when considering the potential effect of an episode-based payment system priced according to average costs on clinical practice, as "small variation around the mean (cost) cannot necessarily be seen to be any better than large variation, if we have no idea what efficient practice is".⁵ A major challenge exists in defining efficient care for service users with multiple problems and complexity factors, which the consultation suggested form a sizeable proportion of cases in practice. It may be productive to explore the feasibility of identifying common treatment components recommended in the guidelines, and recombining them theoretically for cases with co-occurring problems.²⁶ Such a process would likely benefit from additional input from clinicians, particularly with regard to obtaining views on how interaction between particular combinations of problems might either potentiate or lessen resource use compared with single problem episodes.¹

Study 3's strength was its examination of real cases and their range of complexity. Modest associations were found between particular service user attributes and resource use, and the

analysis suggests substantial variation in resource use among CAMHS that cannot be explained with available data. This preliminary finding is consistent with the results of other case-mix classification studies in mental health that analysed out-patient data.^{3, 18, 19} Aside from random error several possible reasons for the inter-service variance can be hypothesised. These include systematic differences between services in (i) unmeasured case-mix, (ii) influence of parents/carers on access to care, (iii) contributions of other providers (e.g. social care, schools), (iv) other external factors (e.g. living conditions), (v) technical efficiency, and (vi) clinical practice. It is also likely that some of the variation is due to differences in recording activity between teams and/or localities. This leads us to propose that a nationally agreed framework of using more detailed and validated assessment tools and costing would enhance the data available for further empirical work.

Notwithstanding the importance of future research to establish whether prediction of resource use can be improved by taking into account complexity and context factors alongside problem types and severity, these preliminary findings suggest that the nature of CAMHS presents a significant challenge to defining episode-based ‘products’ homogeneous enough for relative costs (and subsequently prices) to reflect the relative needs of service users. Even where case-mix classifications arguably deliver better prediction of resource use, as in in-patient care for physical health problems, the extent of within-group variation has raised concern about potential detrimental effects on equity of access and provider financial stability.^{9, 16}

Gradual implementation, comprehensive monitoring of processes and outcomes, and refinement are, therefore, likely to be crucial to CAMHS payment reform.^{4, 9} This will be vital for managing the complexities of the desirable and undesirable incentives that

accompany an episode-based payment system, including consideration of repeat episodes of care in conditions such as depression, when chronicity is predicted by repeat remission.⁶ It is worth noting the potential relevance of the English NHS context here. The introduction of a national system of case-based reimbursement for hospitals in the mid-2000s placed local payers at considerable financial risk for increases in activity. Providers' understanding of the fixed budget for their local health economy and concern for not harming the financial condition of other organisations was suggested to have helped constrain activity growth.¹²

Acknowledgements

This project is part of the development of a payment system for child mental health services, which is funded by the Department of Health and NHS England. The authors' views do not necessarily reflect those of the policy department. We are grateful to the CAMHS Payment System Project Group, the Expert Advisory Group, the participating services, and a range of stakeholders for their valuable contributions.

References

1. Hornbrook M, Hurtado A and Johnson R. Health Care Episodes: Definition, measurement and use. *Med Care Rev* 1985;42(2):163-218.
2. Mechanic R. Opportunities and challenges for episode-based payment. *N Engl J Med* 2011;365(9):777-9.
3. Gaines P, Bower A, Buckingham B, et al. *New Zealand Mental Health Classification and Outcomes Study: Final Report*. Auckland: Health Research Council of New Zealand, 2003.
4. Mason A, Goddard M, Myers L, et al. Navigating uncharted waters? How international experience can inform the funding of mental health care in England. *J Mental Health* 2011;20:234-48.
5. Donaldson C, Gerard K, Jan S, et al. *Economics of Health Care Financing - The Visible Hand*. Basingstoke: Palgrave Macmillan, 2005.
6. Cots F, Chiarello P, Salvador X, et al. DRG-based hospital payment: Intended and unintended consequences. In: Busse R, Geissler A, Quentin W, et al. (eds.) *Diagnosis-Related Groups in Europe: Moving towards Transparency, Efficiency and Quality in Hospitals*. Maidenhead: Open University Press, 2011.
7. Donaldson C and Magnussen J. DRGs: The road to hospital efficiency. *Health Policy* 1992;21:47-64.
8. O'Reilly J, Busse R, Häkkinen U, et al. Paying for hospital care: The experience with implementing activity-based funding in five European countries. *Health Econ Policy Law* 2012;7(1):73-101.
9. Allen P. 'Payment by Results' in the English NHS: the continuing challenges. *Public Money & Management* 2009;29(3):161-6.
10. Fetter R, Shin Y, Freeman J, et al. Case-mix definition by diagnosis-related groups. *Med Care* 1980;18(2 Suppl):1-53.
11. McMahon L. Diagnosis-related group prospective payment: Effects on medical quality assurance. *Eval Health Prof* 1984;7(1):25-41.
12. Farrar S, Sussex J, Yi D, et al. *National Evaluation of Payment by Results. Report to the Department of Health*. Health Economics Research Unit, 2007.
13. Vladeck B and Kramer P. Case-mix measures: DRGs and alternatives. *Annu Rev Public Health* 1988;9:333-59.
14. Bevan G and Price C. Roles of Case-mix measures in managing use of resources. In: Leidl R, Potthoff P and Schwefel D (eds.) *European Approaches to Patient Classification Systems: Methods and Applications Based on Disease Severity, Resource Needs, and Consequences*. Berlin: Springer-Verlag, 1990.

15. Serdén L and Heurgren M. Sweden: The history, development and current use of DRGs. In: Busse R, Geissler A, Quentin W, et al. (eds.) *Diagnosis-Related Groups in Europe: Moving towards Transparency, Efficiency and Quality in Hospitals*. Maidenhead: Open University Press, 2011.
16. Or Z, Häkkinen U. DRGs and quality: For better or worse? In: Busse R, Geissler A, Quentin W, et al. (eds.) *Diagnosis-Related Groups in Europe: Moving towards Transparency, Efficiency and Quality in Hospitals*. Maidenhead: Open University Press, 2011.
17. Taroni F. Using diagnosis-related groups for performance evaluation of hospital care. In: Leidl R, Potthoff P and Schwefel D (eds.) *European Approaches to Patient Classification Systems: Methods and Applications Based on Disease Severity, Resource Needs, and Consequences*. Berlin: Springer-Verlag, 1990.
18. Buckingham B, Burgess P, Solomon S, et al. *Developing a Case-Mix Classification for Mental Health Services*. Canberra: Commonwealth Department of Health and Family Services, 1998.
19. NHS Information Centre. *End Stage Report: Mental Health Case-Mix Classification Development*. Leeds: Health and Social Care Information Centre, 2006.
20. Department of Health. *Project Initiation Document: Mental Health Payment by Results Development Project*. London: Department of Health, 2008.
21. Marshall M. How Should we measure need? Concept and practice in the development of a standardized assessment schedule. *Philosophy, Psychiatry, & Psychology* 1994;1(1):27-36.
22. Rigby A. *Cluster Development Story*. Care Pathways and Packages Project, 2013.
23. Department of Health. *Equity and Excellence: Liberating the NHS*. London: The Stationery Office, 2010.
24. NHS Benchmarking Network. *CAMHS Benchmarking Report, December 2013*. Benchmarking Network, 2013.
25. Fittall, B. *Developing a Payment system for Child and Adolescent Mental Health Services in England*. Presentation at: CAMHS Payment System Project October 2013 Events. <http://pbrcamhs.org/wp-content/uploads/2013/11/CAMHS-data-events-October-2013-v2.pdf> (accessed 10/1/14).
26. Weisz J, Chorpita B, Palinkas L, et al. Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth. *Arch Gen Psychiatry* 2012;69(3):274-82.
27. Deighton J, Croudace T, Fonagy P, et al. Measuring mental health and well-being outcomes for children and adolescents to inform practice and policy: A review of child self-report measures. *Child Adolesc Psychiatry Ment Health* 2014;8:14.

28. National Institute for Health and Care Excellence: Guidance List.

<https://www.nice.org.uk/guidance>

(accessed 10/1/15)

29. Shaffer D, Gould M, Brasic J, et al. A Children's Global Assessment Scale (CGAS). *Arch Gen Psychiatry* 1983;40:1228-31.

30. Child Outcomes Research Consortium (CORC). *Getting started: Snapshot Approach*.

<http://www.corc.uk.net/joining/getting-started/getting-started-snapshot-approach/> (accessed 12/11/14).

Table 1

**Association between service variables and number of appointments attended
(mixed negative binomial regression)**

	Coefficient	Standard Error	Z	p
Intercept	2.89	0.32		
Sex (ref: male)				
female	0.16	0.05	3.00	0.003
Age (ref: 0-5 yrs)				
6-12 yrs	0.27	0.10	2.65	0.008
13-18 yrs	0.23	0.10	2.19	0.028
Problem type (ref: low resource use)				
Moderate/Conduct	0.18	0.11	1.61	0.107
Moderate/Emotional	0.36	0.06	6.40	<0.001
High use	0.73	0.17	4.24	<0.001
CGAS score	-0.0237	0.0019	-12.72	<0.001

Notes:

The log-link was used for the generalised linear mixed model.

Random effect variances: Region: 0.1996; Team: 0.2016.

Negative Binomial dispersion parameter (NB2 parameterisation): 1.0921 (SE: 0.0455).

The age groups followed usual criteria of pre-, primary and secondary school age years.

Table 2

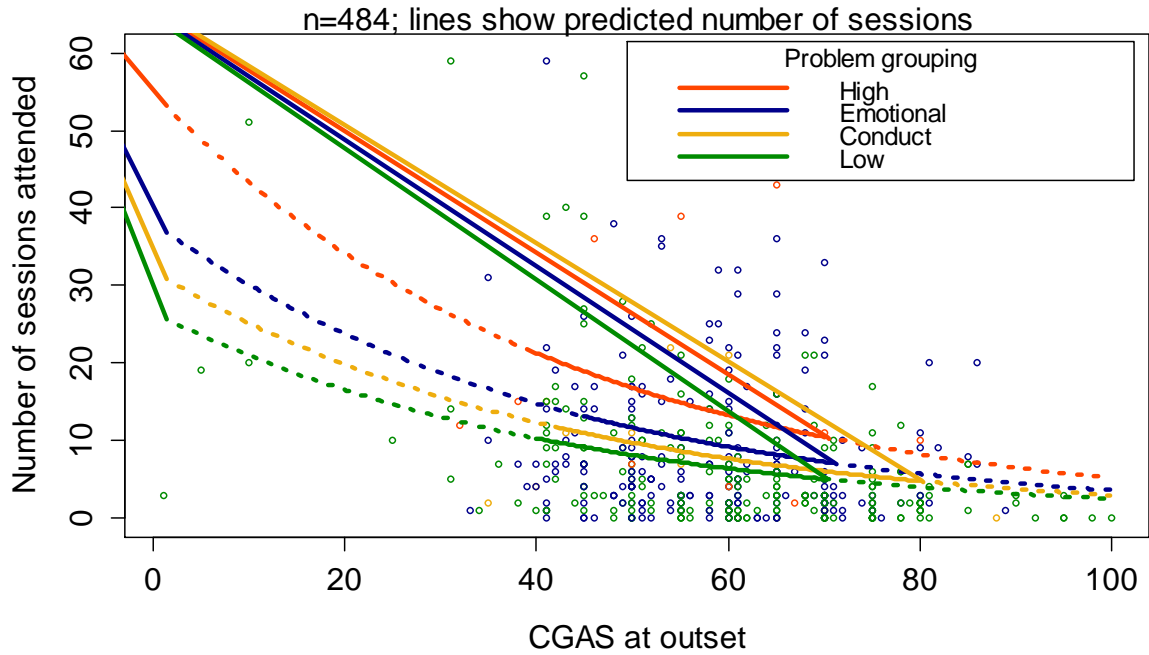
Key findings and methodological issues of studies 1-3

	Key findings	Methodological issues of each study
Study 1: NICE Guidelines Review	<ul style="list-style-type: none"> • Included majority of child mental health problems • Predominantly diagnostic-driven 	<ul style="list-style-type: none"> • Largely not contextualised (i.e. in relation to service implementation, complexity, cost of interventions) • No indirect activities • Evidence complemented by expert consensus
Study 2: CAMHS Staff Consultation	<ul style="list-style-type: none"> • Needs- rather than diagnostic-led • Should not drive clinical decisions • Consider complexity • Measure indirect activities • Link with outcomes and resource use 	<ul style="list-style-type: none"> • National but small sample • Not necessarily representative • Not matched with service data
Study 3: CAMHS Case Study Data Analysis	<ul style="list-style-type: none"> • Service variation in resource use • Psychosis and eating problems associated with high resource use • Less clear pattern for other child mental health problems • Severity and impairment independently predict resource use 	<ul style="list-style-type: none"> • Potential service bias • Measures not selected for this purpose • Inconclusive on treatment outcomes

Figure 1

Appointments by impairment (CGAS score) and problem type (teenage girls)

Treatment sessions by CGAS & problem type (girls aged 13-18)



Note: The solid lines show the middle 80% of cases with respect to CGAS ratings for each problem type. Dotted lines show the bottom and top 10 % of cases. Six teenage girls attended more than 60 appointments. They are excluded from this graph, but were included in the statistical model.