Piloting a new Child and Adolescent Risk Assessment Suite in the United Kingdom

Paul A. Tiffin[1], Charlotte E. W. Kitchen[2] & Steven Weir[3]

[1]Clinical Senior Lecturer, School for Medicine, Pharmacy and Health, Durham University, Queen's Campus, Stockton-on-Tees, TS17 6BH, UK; Telephone: 0191 33 40707; Email: p.a.tiffin@dur.ac.uk

Corresponding author and guarantor

[2]PhD Student, School for Medicine, Pharmacy and Health, Durham University, Queen's Campus, Stockton-on-Tees, TS17 6BH, UK; Telephone: 0191 33 40455; Email: charlotte.kitchen@dur.ac.uk

[3]Consultant to FACE Recording & Measurement Systems Ltd, Nottingham, NG1 7AR, UK; Telephone: 0845 8621259; Email: steve@phoenixsocialwork.co.uk

Abbreviated title: **FACE Child and Adolescent Risk Assessment Suite**

Abstract

Background: A prototype risk assessment suite (FACE-CARAS) was developed for use within CAMHS and evaluated for acceptability and reliability.

Method: Clinicians underwent brief training in the system and invited 69 young people to an assessment using the FACE-CARAS. A second rater produced a separate set of blind ratings for most patients. Clinicians also provided qualitative feedback.

Results: The component schedules of the FACE-CARAS could be reliably rated with 'near perfect' to 'moderate' agreement observed. Internal reliability-consistency values, as indexed by Cronbach's alpha, were moderate to high in all cases.

Conclusions: The assessment schedules that make up the FACE-CARAS can be reliably rated by clinicians with minimal training.

Keywords: risk assessment; adolescents; children

Word Count: 2896

Young people affected by mental health issues frequently present with indicators of increased risk to either themselves or others (McArthur Foundation Research Network on Mental Health and the Law, 1996; Monahan et al., 2000; Tiffin & Kaplan, 2004). Despite this scenario, a comprehensive risk assessment system for use in Child and Adolescent Mental Health Services (CAMHS) is not currently available (Borum, Bartel, & Forth, 2003; Tiffin & Nadkarni, 2010; Tiffin & Richardson, 2006). Existing risk assessment protocols only assess specific risk domains (e.g. interpersonal violence), have been generally developed outwith the UK (Borum et al., 2003; Forth, Kossen, & Hare, 2003), and tend to focus on adults (Cooper & Tiffin, 2006). Risk assessment schedules for young people need to account for the developmental context and dynamic nature of risk factors. Moreover, risk assessment instruments developed in countries with differing levels of baseline violence may not be valid (Singh, Fazel, Gueorguieva, & Buchanan, 2014).

The FACE- Child and Adolescent Risk Assessment Suite (FACE-CARAS) consists of nine novel schedules (no schedules were previously licenced or copyrighted) designed to address this gap. The suite contains a general 'Risk Profile' that can be used in conjunction with any of the nine individual schedules; if required, more than one schedule can be used per patient. Six of these schedules focus on specific risk domains (self-harm, aggression, aggression in psychosis, vulnerability, learning disability vulnerability and sexually harmful behaviour) and the remaining three schedules focus on specific inpatient settings (eating disorders, low secure and open wards). The information collected from these schedules is then summarised and used to inform a risk formulation and risk management plan. The content of FACE Eating Disorder Schedule (FEDS) was informed by the MARSIPAN Junior guidelines (Royal College of Psychiatrists, 2014).

After the initial development of the schedules, focus groups were held with CAMHS clinicians to provide in-depth feedback on the draft FACE-CARAS (Daniel, Weir, & Tiffin, 2013). This feedback suggested that the FACE-CARAS was a generally well-structured and clinically acceptable risk assessment suite (Daniel et al., 2013). The findings were subsequently fed back into the design of the system. The current study aimed to assess whether this amended version of the FACE-CARAS was a practical and feasible approach to risk assessment in CAMHS. Reliability is a necessary, but not sufficient, prerequisite for validity. Therefore, we evaluated the reliability of the FACE-CARAS.

Methods

A mixed quantitative/qualitative methodology was used. The study was conducted within Tees, Esk and Wear Valleys (TEWV) NHS Foundation Trust. A favourable ethical opinion was obtained from the NRES Northern & Yorkshire and the School of Medicine, Pharmacy and Health Ethics Sub-Committee (REC reference: 11NE/0248).

Clinicians were provided with an hour-long training session, including practice vignettes. The FACE-CARAS is completed according to structured professional judgement based upon the responses to an initial screen in the Risk Profile (Figure 1). Therefore, not all schedules were completed for each participant. Each schedule item is rated using between two and five anchor points. For example, in the Checklist for Risk of Aggression in Youth (CRAY) one question on '*Frequency of previous aggression*' has four possible anchor points: '*No problem*', '*Some aggression but less than 4 physically aggressive episodes directed at people per year*', '*Physical aggression towards others occurs on average between 4/year to less than once a month*' or '*Physical aggression towards others occurs at least on average*

monthly'. In this pilot study, the FACE-CARAS was implemented using a paper format.

Participants were recruited via two methods; an internal audit (using the Sexually Harming Adolescent Risk Protocol- 20 Item Version [SHARP20]) or a cross-sectional survey using a purposive sampling approach. Eight individuals were approached for the internal audit and 64 patients who currently required risk assessments (initial, review or discharge) were invited to participate. The sample was derived from seven CAMHS teams (forensic, Early Intervention Psychosis [EIP], a low secure inpatient, open inpatient, Eating Disorder Unit, and two Tier 3 CAMHS. The Learning Disabilities Service [LD] did not receive any qualifying young people in the short time frame of the study. Clinicians, at least one from each team, were approached via their team managers and asked to use the FACE-CARAS. In some assessments, a second rater (either a researcher or another clinician) was present who completed risk ratings blind to the main clinicians coding.

Participant inclusion criteria:

- an active patient requiring a risk assessment as Trust policy

-aged 10 to 18 years of age

-with sufficient English to respond to the questions in the schedules

-who consents to participate (or, if under 16, assents with parental consent)

INSERT FIGURE 1

Figure 1. Layout of the FACE CARAS

Clinicians were asked for feedback on the FACE-CARAS during the study and completed an open-ended questionnaire at completion.

Data Analysis

Analysis was conducted on paired ratings. Inter-rater reliability was reported using the average quadratically weighted kappa for each schedule and internal reliability using Cronbach's alpha. A quadratically weighted kappa was utilised to disproportionately penalise disagreements that differed by more than one point, reflecting the potential clinical implications of disparities in ratings. Cronbach's alpha was calculated on average paired ratings.

Results

Recruitment

A total of 69 participants were approached to take part (six were excluded for incomplete consent [one refusal]) leaving a total of 63 participants suitable for analysis. The mean age of the sample was 15.94 (range 12.23- 18.71) with 36 males; two participant's age and sex were not known due to mis-recorded unique identifiers. Data were obtained from 20 participating clinicians and the research assistant.

Schedules

Risk Profile

Forty nine were completed (44 paired ratings) from all services except Eating Disorders. There were two elements (*self-neglect/accidental self-harm*) of this schedule on which raters (n=2) disagreed by more than one point.

The screening checklist was praised as useful for inexperienced members of staff, however it was also suggested it would be helpful if more experienced staff could bypass this section. The kappa and alpha values for the FACE-CARAS schedules are provided in Table 1.

Child & Adolescent Self Harm Schedule (CASH)

Fourteen were obtained (12 paired) from all services except Eating Disorders and low secure. On elements relating to *substance use, non-life threatening self-harm, low mood/dysphoria* and *problem solving/ability to cope with stress*, one set of raters disagreed by two points. Two sets of raters disagreed by two points on the *suicidal ideation* question.

CRAY

Twenty seven were obtained (22 paired) from forensics, open inpatient and two community teams. Questions relating to *current weapon use*, *frequency and intensity of aggression, quality of care*, *peer relationships, empathy, social competency, sibling criminal history* and *educational exclusions* each had one set of raters disagree by two points. The items relating to *previous weapon use, exposure to domestic violence* and *impulsivity* had two sets of raters disagree by two points and one set by three points. Both items relating to *fire-setting* and *substance use associated with aggression* had one set of raters disagree by three points. The total of the averaged paired ratings scores for the CRAY items was significantly higher for the forensic compared to the non-forensic CAMHS patients (χ^2=19.36, p<0.001 for inter-group difference on Wilcoxon ranked-sum test). This difference was not accounted for by trends in age or

sex as the two groups with completed CRAY schedules did not significantly differ in these respects.

FEDS

Ten were completed (all paired) by the Eating Disorders team. No items demonstrated marked disagreement.

Learning Disability Vulnerability Assessment Schedule (LDVAS)

Two were obtained (all paired) from EIP. The majority of the questions demonstrated 100% agreement between raters. On items relating to *physical impairments/unrecognised sources of pain* and *parent/carer antisocial traits/behaviour* one set of raters disagreed by two points.

Ward Assessment (OPEN)

Eight were obtained (all paired). No raters disagreed by more than one point.

Schedule for Risk of Aggression in Psychosis (ScRAP)

Six were obtained (5 paired) from forensic, community and EIP teams. Items relating to *pre-morbid antisocial personality traits* and *non-concordance* had one set of raters disagree by two points. Questions relating to *treatment concordance, pre-morbid aggression to persons, previous non-concordance, passivity experiences, identified triggers to aggression* and *impaired insight* had two sets of raters that disagreed by two points.

Ward Assessment (SECURE)

Seven were obtained (all paired) from the secure ward. For the item on *absconsion risk*, two sets of raters disagreed by two points.

SHARP20

Ten were obtained; two from the pilot, eight from the internal audit (7 paired) all from forensics. Items relating to *attitudes towards sexually harmful behaviour*, *sexual development status (pre-pubertal)* and *sexual interests and sexual preferences*, one set of raters disagreed by two points. On the item relating to the *nature of aggression* two sets of raters disagreed by two points.

Vulnerability Assessment Schedule (VAS)

Nine were obtained (8 paired) from community teams and forensics. On items relating to *history of abuse or neglect*, *geographical mobility* and *parent/carer antisocial traits/behaviour*, one set of raters disagreed by two points. On items relating to *exposure to domestic violence* and *parent/carer physical health problems*, two sets of raters disagreed by two points.

INSERT TABLE 1

Table 1: Average quadratically weighted kappa and Cronbach's alpha values for the FACE-CARAS schedules.

*** Almost perfect agreement, ** Substantial agreement, * Moderate agreement

Clinician Feedback

Mid-point clinician feedback highlighted inconsistencies between anchor point definitions and their corresponding ratings. All such points were corrected in an iteration following the completion of this project. At the end of the study, eleven clinicians provided further feedback via a questionnaire.

When asked to describe the overall experience of using the FACE-CARAS clinicians described it as "*straightforward*" (n=2), "*functional*", "*good*" (n=2), "*fit well with the client group*" "*excellent learning experience*" "*easier once practiced a few times*", "*satisfactory/good means of gathering info*", "*useful*" "*liked the structure*", "*positive*", "*very comprehensive but flexible system*". Some noted "*the initial assessment form is slightly overpopulated with data*", at times it was "*lengthy and monotonous*" although it was acknowledged that some repetition may be avoided via an electronic format. Clinicians reported they would be content to use the system in the future but some highlighted they "*would be happy to use in clinical practice but not in paper form*" and "*will need to be incorporated into the [Trust electronic records] system*". When asked if they experienced any problems using the FACE-CARAS most clinicians did not report any problems. However, some responded that the initial FACE-CARAS initial screening schedule could be "*streamlined more*", "*time, however, merely a teething problem*", "*some of the items could have been weighted differently to denote different levels of risk*", "*feel the initial screening questions section too lengthy*". When asked what impressed them the most about the FACE-CARAS responses included "*gives a clear/overall picture of clinical risk without collecting too much information*", "*breadth and depth of information covered*", "*clear anchor points*", "*ease of use*", "*comprehensive nature of the tool*", "*very in-depth assessment once completed*", "*comprehensive, helped lead risk-based discussion*", "*how it encourages a thorough consideration of all aspects of a patients*

situation", "*objectivity*", "*comprehensive with client group*", "*clinically relevant*" and "*it was clearly and logically set out*". All participating staff agreed that electronic implementation of the system would be likely to quicken assessment and avoid the need for duplicate entries. One concern raised was that the pilot was only being trialled in children aged 10 and above. One clinician also commented that the schedules tend to be quite "negative" in their language rather than focusing on protective factors.

Discussion

Overall, the instruments that make up the FACE-CARAS were reliably rated by clinicians with minimal training. The high reliability is probably partly due to the clear anchor points provided. A small number of paired ratings showed poor agreement. This may have been due to the complexity of the clinical presentation or erratic rater behaviour. Further research would be required to elicit the cause of non-concordant ratings. The qualitative feedback supports earlier work (Daniel et al., 2013) indicating that the system, even in paper format, was acceptable, although there were some suggestions for improvement. Internal reliability-consistency was generally high, as indexed by Cronbach's alphas. The inter-rater reliability was generally comparable with those for previously trialled risk instruments (Vincent, Guy, Fusco, & Gershenson, 2012). However, whilst high alpha values tend to suggest items are tapping into the same construct (unidimensionality) they may also hint that certain items may be redundant or dependant on responses to other questions. Thus, future studies may highlight where the schedules can be shortened with no loss of information.

The FEDS and LDVAS schedules demonstrated almost perfect agreement between raters. In the FEDS, this may be accounted for, to some extent, by the majority of the questions only having two rating options compared to the other schedules that for the most part use a 4-point rating scale combined with the majority of questions being objective (i.e. temperature) rather than subjective (i.e. therapeutic engagement). Although we have limited data on the LDVAS, preliminary data suggests high inter-rater reliability. This would need to be confirmed by further exploration in an LD population. Forensic CAMHS patients received higher ratings on the CRAY violence risk schedules compared to generic CAMHS, suggesting some discriminative validity.

The OPEN schedule demonstrated low variance in rater responses, highlighting the presence of 'floor' effects. This is a test targeting issue in that many in-patients did not show evidence of risk factors at the times the ratings were conducted. Thus, limited information about participants scoring at the lower end of the rating scales was available. However, this may be appropriate in this setting as in-patient clinical staff should only be alerted to factors which are likely to significantly increase the risks to the patient and others. In contrast, the raters of the SECURE schedule often produced 'non-zero' ratings for their clients suggesting better test targeting due to the higher level of perceived risks. The risk profiles on the SECURE do appear to be potentially able to discriminate between individuals, given the reliable scoring and the reasonable spread of item scores, which is likely to equate to acceptable levels of test information. There may have been some selection bias due to clinicians selecting patients they deemed more risky, which may have increased the reliability of the codings. The suite may be less useful for 'low risk' cases. Only one patient refused participation which may have reduced the risk of response, if not

selection bias. Clinician training for the FACE-CARAS should focus on items that demonstrated a low kappa score or they should be removed from future iterations of the suite.

The main limitation of the study was that the FACE-CARAS could not be implemented in its intended electronic format. An electronic 'Beta' version of the system has now been produced and is undergoing post-marketing testing. One concern raised was that the pilot was only being trialled in children aged 10 and above. If the FACE-CARAS is to be used in practice in younger children it is important that it is validated in this population which was not part of the remit of this study. We did not explore the impact of clinical experience on reliability of ratings in this study; although there were no indications from our small sample that experience effected reliability. The number of pairings was small for some schedules and although this is acceptable for a pilot study, it limits generalizability. Our relatively small sample from one UK-based mental health Trust may not generalise to other healthcare settings. Ideally work streams should also be developed that will gather evidence to support or refute the concurrent, and ultimately, the predictive validity of the individual tools that make up the risk assessment system.

References

Borum, R., Bartel, P., & Forth, A. (2003). *Manual for the structured assessment of violence risk in youth: Version 1.1- Consultation edition*. Florida: University of South Florida.

Cooper, S., & Tiffin, P. A. (2006). Psychological assessment and treatment of adolescent offenders with psychopathic personality traits. *Educational & Child Psychology, 23*(2), 6274.

Daniel, M. R., Weir, S., & Tiffin, P. A. (2013). A novel CAMHS risk assessment system: clinicians' views. *Journal of Forensic Practice, 15*(3), 182-191.

Forth, A. E., Kossen, D., & Hare, R. D. (2003). *The Hare Psychopathy Checklist - Youth Version (PCL:YV)*. New York: MultiHealth Systems Inc.

McArthur Foundation Research Network on Mental Health and the Law (1996). *Violence, competence and coercion: Pivotal issues in mental health law*. Paper given at Wadham College, Oxford, 4th July, 1996.

Monahan, J., Steadman, H. J., Applelaum, P. S., Robbins, P. C., Mulvey, E. P., Silver, E., Roth, L. H., & Grisso, T. (2000). Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry, 176*, 312-319.

Royal College of Psychiatrists. (2014). CR168. Junior MARSIPAN: Management of Really Sick Patients under 18 with Anorexia Nervosa. Retrieved 28/05/2014, 2014, from http://www.rcpsych.ac.uk/usefulresources/publications/collegereports/cr/cr168.aspx

Singh, J. P., Fazel, S., Gueorguieva, R., & Buchanan, A. (2014). Rates of violence in patients classified as high risk by structured risk assessment instruments. *The British Journal of Psychiatry, 204*, 180-187.

Tiffin, P. A., & Kaplan, C. (2004). Dangerous Children: Assessment and Management of Risk. *Child and Adolescent Mental Health, 9*(2), 56-64.

Tiffin, P. A., & Nadkarni, J. (2010). Evaluating Violence Risk in Young People. In E. Campbell & J. Brown (Eds.), *The Cambridge Handbook of Forensic Psychology*. London: Cambridge University Press.

Tiffin, P. A., & Richardson, G. (2006). The use of structured instruments in assessing risk of violence. In A. Hagell, R. Dent & J. Kingsley (Eds.), *Children who commit Acts of Serious Interpersonal Violence Messages for Practice*. London: Jessica Kingsley Publishers.

Vincent, G. M., Guy, L. S., Fusco, S. L., & Gershenson, B. G. (2012). Field Reliability of the SAVRY With Juvenile Probation Officers: Implications for Training. *Law and Human Behavior, 36*(3), 225-236.

Acknowledgements

We would like to thank all the young people and clinicians who participated in this study. Special thanks go to Dr Dasha Nicolls (Great Ormond Street Hospital) for allowing us to base the FEDS on the MARSIPAN Junior Guidelines and Mr Graeme Richardson (Northern Forensic Service for Young People) for allowing us to base the SHARP20 on an earlier version of the SHARP schedule that he authored.

Funding: This project was funded by a Knowledge Transfer Partnership Grant (KTP007657) with contributions from both the FACE Measurement and Recording Co. and Tees, Esk and Wear Valleys NHS Foundation Trust.

Key Practitioner Message:

- A flexible system of structured clinical risk assessment tools was acceptable for routine use in a variety of CAMHS settings.
- The individual schedules that made up the risk assessment system demonstrated acceptable to high levels of inter-rater reliability and internal reliability-consistency.
- Minimal training in the system (one hour orientation session) was required to achieve inter-rater reliability.

Table 1

Schedule	No. of paired ratings	No. items in the scale	Median kappa value for ratings	Inter-Quartile Range for kappas	Minimum kappa for a patient rating	Maximum kappa for a patient rating	Mean quadratically weighted kappa for patient ratings	Internal reliability-consistency (Cronbach's alpha
Risk Profile	44	8	0.73	0.59	0.00	1.00	0.66**	0.77
CASH	12	17	0.73	0.24	0.00	1.00	0.67**	0.90
CRAY	22	28	0.71	0.38	0.00	0.93	0.66**	0.92
FEDS	10	26	0.98	0.05	0.79	1.00	0.96***	0.73
LDVAS	2	20	0.82	0.35	0.65	1.00	0.82***	0.98
OPEN	8	7	0.53	0.49	0.00	0.89	0.44*	0.68
ScRAP	5	23	0.76	0.41	0.00	1.00	0.60*	0.89
SECURE	7	9	0.74	0.21	0.29	1.00	0.70**	0.81
SHARP20	7	28	0.75	0.36	0.17	0.85	0.65**	0.96
VAS	8	20	0.72	0.45	0.00	0.88	0.59*	0.91