DJI
Deutsches
Jugendinstitut
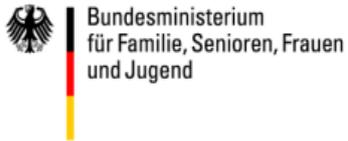
A LITERATURE REVIEW ON SYSTEMS FOR EARLY PREDICTION AND RISK DETECTION IN CHILD PROTECTION IN EUROPE

Eine Expertise im Auftrag des Informationszentrums Kindesmisshandlung / Kindesvernachlässigung

Prof. Kevin Browne, Ph.D.
Shihning Chou, M.Sc.
School of Psychology, University of Liverpool
WHO Collaborating Centre on Child Care and Protection, Birmingham

Wissenschaftliche Texte

Background

Poor parenting, dysfunctional parent child interactions and deprived family environment are associated with child maltreatment and subsequent poor mental health and behavioural outcomes in children (Browne & Herbert, 1997; Farrington, 1995). Therefore, improvements in parenting and the family environment may prevent child maltreatment and future adverse outcomes in children. Research has shown promising developments in early interventions which address parental competency and psychological state and promote positive knowledge, attitudes, skills and behaviour (e.g. Barlow, Coren, & Stewart-Brown, 2003; Coren & Barlow, 2001; Elkan et al., 2000; Kendrick, Barlow, Hampshire, Polnay, & Stewart, 2007).

Services provided to children and their families generally fall into three categories: universal, targeted and specialist. Universal services are offered to all children and families (primary prevention). For instance, health visitors in the UK provide support and advice on parenting skills, child health and parental well-beings to families with all children less than five years old through ongoing visits. This system has been long established since the 19th century. Targeted services (secondary prevention) are aimed at families with identified needs. An example of this is the Nurse Family Partnership (NFP) Programme developed by Olds and his colleagues in the US where low income young mothers with no previous live births are offered intensive home visiting (Olds, Henderson, Jr., Chamberlin, & Tatelbaum, 1986). Specialist services (tertiary prevention) are offered to cases where maltreatment has already occurred. The Physical Abuse Treatment Outcome Project was designed for this purpose (Chaffin et al., 2004), as intervention are offered to parents who have already maltreated their children and the aim is to reduce the likelihood of the reoccurrence of child maltreatment. However, tertiary prevention has been shown to be ineffective as one in four abused children in the UK referred to child protection police units were re-referred for another abusive incident within 27 months, even after social work intervention (Hamilton & Browne, 1999).

These three tiers of services are not necessarily mutually exclusive. From a preventative point of view, more intensive support can be targeted at families with greater needs and provided in addition to a universal provision (Browne, Douglas, Hamilton-Giachritsis, & Hegarty, 2006; Elkan, Robinson, Williams, & Blair, 2001). In fact, early interventions are most likely to succeed if effective and practical screening assessments are in place to ensure that those who need more support are appropriately targeted (Cox, 1997).

Principles of screening and early prediction tools

Assessment instruments used to predict child maltreatment by health and social service professionals are retrospectively developed form a comparison of demographic, individual and psychosocial factors associated with those

who are known to have maltreated their children compared with those who have not maltreated their children. A checklist of risk (and sometimes protective) factors that show significant differences between these two groups are then constructed and applied prospectively to a sample or population. The principle is that the more risk factors present (or a certain combination of risk factors), the higher the probability that child maltreatment may occur in the future. To simplify screening procedures, a cut-off score is generally used in order to classify families with children into high or low risk. To determine whether the classification has been successful, data on outcome must be collected as to whether there is sufficient evidence that a child is eventually maltreated. A 'gold' (or reference) standard measure is used to validate the identification of the maltreatment, for example, the child being placed on child protection register (soft evidence), medical practitioners identifying suspected or actual non-accidental injuries (medium evidence) or the parents being convicted in court for child maltreatment beyond all reasonable doubt (hard evidence). It is important that the results of the reference test are very close to truth. Without the use of a reference standard measure, the performance of the instrument under development will be poorly estimated. In order to make sure the accuracy of a reference measure, efforts are often made to involve several sources of information. The length of follow up of participants is also important with long periods of follow up being more accurate.

To establish the performance of an assessment tool in identifying those with or without a condition, we look at the proportion of those who have a condition being correctly identified by the assessment (sensitivity) and the proportion of those without a condition being correctly ruled out (specificity). For practitioners, it is useful to look at the proportion of those tested positive really having the condition as confirmed by reference test (positive predictive value, PPV) and the proportion of those tested negative actually not having the condition (negative predictive value, NPV). Predictive values, in other words, represent the probability of a test result being correct. Although it is more useful for practitioners to look at predictive values, they are more susceptible to variations in the prevalence of the condition/outcome in the targeted population or sample. Sensitivity and specificity, on the other hand, are more resistant to changes in prevalence and only change when the cut-off point changes or when the population the assessment applied to is dramatically different from the sample tested during development and standardisation. Nevertheless, the more rare the condition in a population, the more difficult it is to develop an accurate screening procedure. This is one of the main obstacles in assessing risk for child maltreatment (Browne et al., 1997).

The most ideal test should identify everyone with the condition or who will develop the condition as positive (100% sensitivity) and everyone without the condition or who will not develop the condition (100% specificity). In reality, it is very unlikely for a test to have a perfect cut-off point which would achieve 100% sensitivity and specificity. There is usually a trade off between sensitivity and specificity, depending on the overall cost of false positives and false negatives, when making a decision on the cut off point.

Box 1 contains a contingency table illustrating the concept and calculation of the above figures.

Box 1.   Contingency table for the relationship between assessment results and true outcomes

		Condition (as determined by reference standard measure)			
		True	False	Total	
Test results (as determined by screening assessment)	Positive	TP	FP	TP+FP (Total tested positive)	PPV = TP/(TP+FP)
	Negative	FN	TN	FN+TN (Total tested negative)	NPV= TN/(FN+TN)
	Total	TP+FN (Total with the condition)	FP+TN (Total without the condition)	Total sample	
		Sensitivity = TP/(TP+FN)	Specificity= TN/(FP+TN)		

TP = true positives

TN = true negatives

FP = false positives

FN = false negatives

PPV = positive predictive value

NPV = negative predictive value

In medical research, it is relatively straightforward to determine whether a disease is present or occurs, as there are usually clearly observable symptoms or signs under physiological and physical examinations. However, it is more difficult when a prediction or identification is to be made on a psychosocial or behavioural problem, as often there are not concrete physical signs but a cluster of psychosocial factors that may involve certain degree of subjective judgements in the identification. More importantly, when the problem behaviour in question is against the law or simply unacceptable under the societal norm, ethical and legal ramifications of the labelling and misidentifications are serious. Therefore, assessments on violent or abusive tendency cannot be carried out before evidence has found to suggest the possibility or confirm the suspicion. The dilemma for the prediction of child maltreatment is whether to increase sensitivity and pick up nearly all the potential maltreating parents (hits) at the cost of high numbers of false positives or alternatively to increase specificity to prevent parents being incorrectly identified as high risk at the expense of potential maltreating parents being missed (false negatives).

The most suitable professionals to assess and intervene

Among all child care and primary health care professionals, community nurses (health visitors and midwives) are often considered most suitable to identify need

and work with vulnerable families before child maltreatment occurs, especially in countries with established universal health care provisions, as they proactively visit families expecting or with newborns (Pinheiro, 2006; World Health Organisation, 2002). Midwives offer individual care to women and their families, helping them take part in their own care planning during pregnancy. In the UK, support from midwives continues from the confirmation of the pregnancy through the post-natal period (not less than 10 days after the end of labour and for longer if the midwife considers necessary). After the post labour visit, the majority of the cases are handed over to health visitors who visit all families with newborns and offer support to parents. These qualified nurses who have undergone further training on child health surveillance as well as the prevention of health problems and injury among all age groups. These two groups of professionals have the unique access to reach families and are generally perceived as playing a caring and supporting role, which put them in a better position to identify need and work with vulnerable families who may be in need of support and advice on parenting and child health (Rowe & Carey, 2004; Rowe, McClelland, & Billingham, 2001). The universal nature of community nursing also has the advantage of NOT stigmatising families unlike intervention by social workers (Rowe et al., 2001).

An example of using midwives for the identification of parents needing more intensive input is the Oxfordshire Intensive Home Visiting Study (Barlow et al., 2007; Barlow, Stewart-Brown, & Callaghan, 2003). It was a randomised controlled trial testing the effectiveness of an intensive parenting programme delivered by health visitors for families in needs. The research team used midwives attached to 40 GP clinics for the assessment of women registered to their clinics on the level of need. A list of nine well established social, demographic and psychological factors associated with poor parenting and child maltreatment were considered (please see Table 3). Midwives carried out the assessment in a consultation during the second trimester and women who met any of the criteria were approached for consent to participate in their trial. The main reason for using midwives was that midwives have the access to pregnant women and are generally seen as in a caring capacity, women are more likely to talk to them and perhaps more amenable to their suggestion about 'having more support' after birth. From a scientific point of view, it helps reduce bias if the recruitment and allocation of participants are concealed from those who actually carried out intervention and research evaluation.

Early assessment instruments used by nurses

In the mid 1970s, the new concept of child abuse and neglect by parents instigated research and the development of assessment tools in hospitals for the early prediction and prevention of child maltreatment. Various instruments were developed for use by health professionals who would come into contact with families expecting and/or with newborns in hospitals. Assessment methods include checklists, questionnaires, structure interviews, behavioural observations and clinical judgements.

In the UK, Gordon (1977) developed a 6 item checklist looking at socioeconomic, demographic and clinical variables for identifying families at high risk of maltreating their children. It was reported to be used by professionals during perinatal period at the hospital but the actual time of assessment was not specified. The outcome measure is only limited to overt cases of battering and the tool achieved 100% sensitivity (all 4 abusers were classified as high risk) and 94.3% specificity. However, the positive predictive value is not promising (1.7%) due to a large number of false positives, which means only 1.7% of the high risk groups actually abuse their children and therefore limits its practical use. Nevertheless, the negative predictive value of 100% means those who are identified as low risk are all true nonabusers.

Another UK study (Lealman, Haigh, Phillips, Stone, & Ord-Smith, 1983) developed a 10 item scale, looking at social, demographic and clinical data, for the assessment of the potential for child maltreatment. The assessment is carried out by maternity unit staff before mothers are discharged. The outcomes measured include both child abuse register held by social services and failure to thrive (FTT). Where information is available, the sensitivity for FTT (97.6%) was better than child abuse (66.1%) but it is the opposite when it comes to specificity (17.9% for FTT and 83.8% for child abuse). Like Gordon (1977), the positive predictive values all prove to be poor (10.7% for FTT and 3.8% for child abuse) while negative predictive values were more useful (99.6% and 97.5% respectively).

An early study in Oxford (Lynch & Roberts, 1978) attempted to use 'early alerting signs' to predict mothers at high risk of child maltreatment in a maternity hospital. Nurses identified the following factors:
1. Mother under 20 years of age with first child
2. Signs of emotional disturbance recorded within the maternity notes
3. Mother referred to hospital social worker
4. Infant referred to special care baby unit
5. Mother shows poor parenting capacity

With a cut-off point of two or more factors, 35 of 50 maltreating mothers were identified but 5 out of 50 non-maltreating mothers were also identified as high risk (false positives). Therefore, the sensitivity was 70% and the specificity was 90%. However, with such a small sample size, it is difficult to determine the performance of these five factors applied to a large population of births. Furthermore, social factors affecting the mothers capacity to parent (eg; social isolation, violent partner) were not considered.

A similar approach was taken in Denver, USA by (Gray, Cutler, Dean, &

Kempe, 1977) used a combination of approaches in the assessment. It includes an interview, a questionnaire and observations of parenting behaviour and parent-child interaction during labour, delivery and the postnatal period from nurses. They concluded that information gained from observers (nurses) in the delivery room was the most accurate in predicting potential for poor parenting (76.5%) compared to prenatal interviews (54.4%) and postnatal observation (54%).

However, regardless of the test accuracy, the implementation of the above assessments is limited to mothers who attend hospitals and give birth in a maternity unit. If solely relying on them, it is most likely to miss out the more vulnerable and hard-to reach individuals in a population and reduces the effects of prevention.

Risk factors

Pregnant mothers at risk of abuse, neglect and abandonment of their forthcoming child are best assessed by nurses in the community as part of a pregnancy surveillance programme. Risk factors associated with high risk pregnant mothers are:

- Unwanted pregnancy
- Poor attendance to routine health check-ups in pregnancy
- Teenage mother under 20 years
- Single or in an unstable relationship
- Poverty and economic problems
- Housing and accommodation problems or homelessness
- Poor mental health and/or a history of mental illness
- Alcohol and drug abuse
- Smoking in pregnancy
- Complications in pregnancy and child birth
- Social and emotional isolation
- History of domestic violence
- Victim of physical and emotional abuse in pregnancy
- Victim of abuse in childhood
- History of being in public care
- Learning difficulties
- Poor educational background or exclusion from school
- History of antisocial behaviour and delinquency
- Poor relationship with parents or in-laws

After birth, a checklist of factors associated with physical abuse and neglect of children was developed into a well known and perhaps sometimes misused psychometric risk assessment is entitled the Child Abuse Potential (CAP) Inventory. The CAP inventory was developed in the USA (Milner, 1989; Milner, 1994; Milner, Charlesworth, Gold, Gold, & Friesen, 1988; Milner, Gold, Ayoub, & Jacewitz, 1984; Milner, Gold, & Wimberley, 1986; Milner & Wimberley, 1979) and has been cross validated in Belgium (Grietens, De Haene, & Uyttebroek, 2007), Croatia (Pecnik & Ajdukovic, 1995) and Greece (Diareme, Tsiantis, & Tsitoura, 1997). It is a 160 item self

report measure, which contains a 77 item physical abuse scale. This scale covers six major factors: distress, rigidity, unhappiness, problems with child and self and problems from others. The measure also contains three validity scales: a lie scale, a random response scale and an inconsistency scale. These three scales are used in various combinations to form three response distortion indexes: faking good, faking bad and random response. Two scales have since been added to the CAP inventory to assess ego strength (Milner et al., 1988) and loneliness (Milner, 1990).

The CAP Inventory has achieved good reliability and validity in general but it should be noted that this instrument was not designed for early identification in primary prevention. Rather, it was designed for assessment in cases where a group of high risk individuals have been identified and the professional would like to quickly screen this identified population for a subgroup of individuals who are most likely to be at risk for child physical abuse. In other words, it can be used in risk assessment during child care proceedings, social work investigations in or evaluations of secondary or tertiary prevention programmes on child physical abuse, although a small proportion of professionals find the tool also useful in spouse abuse (17%) and elder abuse (12%) (Milner, 1989).

Assessments by community nurses in Europe

In Europe, there have also been attempts to develop early prediction instruments to be applied after birth. Agathonos-Georgopoulou and Browne (1997) developed a weighted 15 item checklist, covering child health, parental health and psychosocial risk factors, in Greece. These 15 risk factors are

High predictor
- Child's bad state of hygiene upon referral
- Parents with mental health problems
- Bad quality of relationship between parents
- Parents with adverse life experiences
- Mother strictly disciplined by own parents

Medium predictors
- Stressful life events for parents in the last year
- Child not with both natural parents
- Mother relies upon nobody when in crisis
- Father with unsteady employment or unemployed
- Mother less than 21 years at birth of the child

Low predictors
- Delayed psychomotor development
- Child had 'other illnesses' prior to referral
- Not breast fed as neonate
- Parents expect immediate obedience from child
- Parental absence prior to referral

Risk factor data was retrospectively collected on an abusing and a non-abusing sample by the first author (psychologist), a paediatrician and social workers. It achieved 93% sensitivity and 93.5% specificity. However, this scale has never been applied prospectively by a specific group of health or social service professionals on a population sample and no specific proposal has been made for its use by nurses. Nevertheless, this tool is still worth mentioning, as this was developed and piloted in Greece at a time of economic transition. Differences in cultures and economic needs may result in different risk factors or the relative importance of risk factors being more prevalent in the Balkans compared to other countries in Europe.

During the 1980s, in the UK, similar structured risk assessment procedures were proposed for use by health visitors to identify potential abusing/neglectful parents with a view to intervene before child maltreatment actually occurs after (e.g. Browne & Saqi, 1988; Johnson, 1985). However, it caused great debate and controversy. Like Gordon (1977) and Lealman et al. (1983), all the screening instruments developed for nurses to use resulted in high number of false positives and false negatives due to the low base rate of child maltreatment in general population (Appleton, 1994; Peters & Barlow, 2003). It is difficult to increase the accuracy because child maltreatment is a rare complex phenomenon which involves a number of causal factors. Ethical concerns over the labelling of parents who have not and perhaps will never abuse their children have been raised. For example, a child injury are more likely to be regarded as *non-accidental* if professionals are aware of them being classed as 'high risk' and this is dangerous when the assessment instrument produces a high number of false positives. Many nursing professionals also worry that the role in screening for child maltreatment can make building a working relationship between them and expectant parents more difficult, as parents may perceive health visitors in a policing rather than supporting capacity and therefore antagonise further intervention and advice offered by health visitors.

The universal approach of health visiting in the UK partly reflects an attempt to avoid stigmatising families but it is simply not practical to offer the same level of support to all families under the limited resources. Besides, while all families are offered basic health visiting services, families with greater needs should get more than the basic support regardless of whether they would actually go on to maltreat their children without intervention. Indeed, the legal basis for providing targeted services to children and their families with greater needs was set out in the UK Children Act 2004, which included guidelines for a Health Needs Assessment Tool to be used by health visitors (seeCowley & Houston, 2003).

Some programmes target families in need using only one or two risk factors such as young mothers who are economically disadvantaged and inexperienced in parenting. For example, the NFP Programme in the USA targeted first time young single mothers with low income (Olds et al., 1997; Olds et al., 1986). They argue that the level of need among parents is not determined by the number of risk factors present but the belief they are in control (Olds et al., 1986). A very similar selection approach was used in the Community Mother's Programme in Ireland (Johnson, Howell, & Molloy, 1993; Molloy, 2002) where support was provided primarily to first time

young mothers in economically deprived areas. The biggest distinction between the two programmes is that the former is carried out in a country where there is no universal (primary) provision of health visiting whereas the latter is offered alongside a long standing health visiting tradition and community support. However, such selection method is crude and misses out many others who also need support.

Recently, the NFP programme has been piloted in England to reduce social exclusion in young mothers and the potential for antisocial and delinquent behaviour in their children in later life. However, Browne (2007) applied the NFP entry criteria first time single mothers with low income to an English population in Essex. Of 409 single mothers less than 21 years, 342 were first time mothers and 67 had more than one child. Within five years, 3.5% (n=12) of the first time mothers were subject to child protection procedures in comparison to 12% (n=8) of those with more than one child. Therefore, if the programme only targets first time mothers, it would miss 40% of those young single mothers who are suspected of maltreating their child. With regard to 603 mothers with economic problems in the same population, 31 mothers were subject to child protection procedures within the first 5 years (5%). Nine of these 31 mothers (29%) were first time mothers and 22 (71%) had more than one child. Therefore if the programme only targets first time mothers with economic problems, it would miss the vast majority of families who maltreat their child and fail to reach the objective of significantly reducing antisocial behaviour and delinquency in children with a history of maltreatment.

Even though the NFP has demonstrated effectiveness in improving parenting and reducing child abuse (Olds et al., 1997; Olds et al., 1986), it can only be applied to the same type of population selected in their trials (Olds, 2007). It may not show the same level of effect when applied to individuals with greater needs or other types of needs such as domestic violence. Indeed, it has been argued that greater effects are shown in programmes targeting one or two factors (like the NFP) because families with a constellation of complex needs may simply take longer and resources to achieve the same level of change (Guterman, 1999). It also remains a question whether the same programme can achieve the same level of effect when being applied to areas already with universal health visiting in place (Browne, 2007).

By contrast to targeting families on one or two factors, the CARE programme (Browne et al., 2006) incorporated a multifactorial needs assessment, the Index of Need. This has been developed from earlier research (Browne, 1995; Browne et al., 1997; Browne et al., 1988). Those families identified as 'high priority' with a score of 5 or more from the 14 screening factors were then entered into the Parent Adviser Partnership (PAP) programme (Davies, Day, & Bidmead, 2002a; Davies, Day, & Bidmead, 2002b). The overall aim is to promote positive parenting and prevent child maltreatment. The intervention is carried out alongside the standard health visiting provision in the UK after families have been identified as 'high priority' and 'low priority' for services by the health visitor. Both the assessment and the intervention were developed in the UK and it is thought to be more suitable to be applied in the UK cultural context. Both the assessment and the intervention have been evaluated in 'the real world' on the UK popula-

tion, demonstrated effectiveness and gain professional acceptance (Barlow et al., 2007; Browne et al., 2006). A similar programme is the European Early Promotion Project whose primary focus is on the prediction and prevention of child mental health issues (Puura et al., 2002; Tsiantis, Smith, Dragona, & Cox, 2000).

Methodological problems with research evidence

Due to the disputes over the accuracy and the ethical issues surrounding the identification of potential child abusers, a lot of assessment tools or procedures developed for this purpose seem to have considered unsuitable for practice implementation on a large scale (e.g. Burns, 1985; Gordon, 1977; Gray et al., 1977; Lealman et al., 1983). In most countries, little attempt has been made to provide strong evidence for the use of screening assessments for child maltreatment. A prospective cohort study with a long enough follow up period is the best research design. However, in countries with universal health visiting, bias to the results may occur because support from health visitors may create positive changes in their parenting, which in turn is likely to reduce the incidents of child maltreatment. A opposite bias is the likely increase in surveillance and attention given to families classified as high risk, which leads to more cases being detected within this group.

However, the need to prioritise health visiting services due to limited resources can never be eliminated. As a result, many professionals resort to using assessments and following guidelines that lack evidence of their effects. A wide range of assessment systems have been found to be used by health visitors in the UK, most of which are not supported by any evidence of effect (Appleton, 1997) and there are huge variations or even contradictions in the understanding and implementation of the risk factors and assessment guidelines (Appleton, 1995; Appleton & Cowley, 2004). Nevertheless, some of the tools developed earlier have continued to transform, within the ethical boundaries and financial limits, into practice based assessments for health visitors to identify the level of need in children and their families. Some have been evaluated with a prospective cohort design (e.g. Browne et al., 2006). There are also new assessments that have been developed in recent years but the follow up data is yet to be seen (e.g. Grietens, Geeraert, & Hellinckx, 2004).

Aims and objectives

The aim of this report is to (1) review published assessment approaches that have been used in public health practice in Europe for early identification of families in need who are at risk of child maltreatment, and (2) to evaluate the accuracy and quality of research evidence behind those assessment procedures that have been developed. As the practices in community nursing

and child protection have changed considerably over the past two decades, the focus is placed upon instruments developed in or after 1985. Studies on the prediction on child maltreatment during the perinatal period prior to 1985 have been reviewed by Leventhal (1988). However, many of these early studies were of poor quality and some exception to this have been included in the background to this review.

Study inclusion Criteria

Population: Children and their families in Europe

Exposure:
a) Screening/risk prediction, developed in Europe in or after 1985, of poor parenting, poor parenting or child maltreatment carried out by community nurses.
b) Assessment procedure, developed in the Europe in or after 1985, to identify the level of need for families in order to prioritise services offered by community nurses.

Methods

Search strategy

In order to identify assessment systems, systematic reviews and meta-analyses on the early prediction instruments for child protection used by community nurses, online databases were searched using the search terms below. Combination operators (AND, OR, NOT) were applied to the primary source citation results.

Search terms used are as follows:

> Child abuse/child neglect/child maltreatment/child protection
> AND
> Risk assessment/risk factors/screening
> AND
> Health visits/home visits/nurse visits

Source of information

Online databases:

The Cochrane Library gateway
Medline (1985 – September 2007)
EMBASE (1985 – September 2007)
CINAHL (1985 – September 2007)
British Nursing Index and Archive (1985 – September 2007)
ISI Web of Knowledge (1985- September 2007)
Applied Social Sciences Index and Abstracts (ASSIA) (1987 – September 2007)

Key literature reviews:

The following literature reviews were identified and their reference lists were hand searched to identify relevant studies.

Appleton, J. (1994). The role of the health visitor in identifying and working with vulnerable families in relation to child protection: a review of the literature. *Journal of Advanced Nursing, 20,* 167-175.

Elkan, R., Kendrick, D., Hewitt, M., Robinson, J. J. A., Tolley, K., Blair, M. et al. (2000). The effectiveness of domiciliary health visiting: A systematic review of international studies and a selective review of the British literature. *Health Technology Assessment. 4,* 1-338. Winchester, England.

Geeraert, L., Van den Noortgate, W., Grietens, H., & Onghena, P. (2004). The effects of early prevention programs for families with young children at risk for physical child abuse and neglect: A meta-analysis. *Child Maltreatment, 9,* 277-291.

Hahn, R. A., Bilukha, O. O., Crosby, A., Fullilove, M. T., Liberman, A., Moscicki, E. K. et al. (2003). First reports evaluating the effectiveness of strategies for preventing violence: early childhood home visitation. Findings from the Task Force on Community Preventive Services. *Morbidity & Mortality Weekly Report. Recommendations & Reports., 52*, 1-9.

Peters, R. & Barlow, J. (2003). Systematic review of instruments designed to predict child maltreatment during the antenatal and postnatal periods. *Child Abuse Review, 12,* 416-439.

Shaw, E., Levitt, C., Wong, S., Kaczorowski, J. & the McMaster University Postpartum Research Group (2006). Systematic review of the literature on postpartum care: effectiveness of postpartum support to improve maternal parenting, mental health, quality of life, and physical health. *Birth, 33*, 210-220.

Quality assessment

All the validation and evaluation studies for the included assessment systems are assessed for quality of study and reporting, using a set of predefined criteria. All the assessment tools are assessed for their theoretical grounding, research base and practical utility, using asset of predefined criteria.

Data extraction

Data from all the included studies were extracted, using a structured form.

Results of search and selection

The search yielded 1351 hits. Of these, 343 of which were duplicates, 973 did not meet the inclusion criteria, 23 were literature reviews or opinion papers and only 12 met the inclusion criteria. Out of these 12 publications, nine assessment packages were identified. However, one of the nine reported few details on the assessment package used and the authors never answered our query after being contacted. Therefore, eight assessment systems are included in the critical review. Figure 1 below illustrates the study selection and sorting process.

Figure 1: Search results and evaluation of primary studies

Total hits = 1351

23 publications = literature reviews or opinion papers of relevance

343 publications = duplicates

973 publications = did not meet inclusion criteria

12 publications meeting inclusion criteria

9 assessment systems identified out of the 12 publications

1 did not provide details on the assessment

8 systems reviewed:
  5 have been statistically validated or evaluated
  3 have not been statistically validated or evaluated

Descriptive data synthesis

Overall characteristics of the assessment systems

Validated assessment tools

Of these eight assessment tools, three have never been standardised nor statistically tested for reliability and/or validity (Ammaniti et al., 2006; Barlow et al., 2003; Cowley et al., 2003) although one of them has been evaluated qualitatively (Cowley & Houston, 2003); five were standardised and validated but only one has been evaluated in its predictive validity and practical application. Table 1 summarises characteristics of five instruments that have been validated and/or evaluated. Table 2 summarises the implementation of the validated tools. Table 3 and 4 summarises characteristics and implementation (respectively) of those instruments that have not validated.

Of the five validated assessment tools, four were developed in the UK (Brockington et al., 2001; Browne et al., 2006; Johnson, 1985) and one was developed in the Flemish community of Belgium (Grietens et al., 2004). The objectives of those tools varies from the identification of general level of psychosocial need in relation to child care and parenting in the family (Browne & Saqi, 1988; Browne et al., 2006), the risk of child maltreatment (Grietens et al., 2004; Johnson, 1985) to a specific problem, bonding disorders (Brockington et al., 2001). Browne et al. (2006) made it clear that Index of Need is NOT to be used alone to identify children likely to suffer significant harm because without behavioural and parenting observation as a second procedure, the Index of Need is likely to produce a large number of false positives. This was the case in the Browne and Saqi (1988) study. More details will be given later in the summary of each system. All of them use health visitors to carry out the assessment although Brockington et al (2001) can also be used by midwives.

All the tools are short in length, ranging from 12 to 25 items, which makes them relatively user friendly. Three of the tools (Browne & Saqi, 1988; Browne et al., 2006; Johnson, 1985) are weighted in a way that those factors that are considered more important are assigned higher scores than the others and all except Grietens et al. (2004) clearly stated the cut off point for a family/mother to be considered high risk/need. However, it should be noted that from the published materials, only the CARE programme based their decisions on the cut-off point on statistical analyses. Johnson (1985) stated that the cut-off for her tool was chosen arbitrarily and no further details were given on the decision making.

In terms of the implementation of those tools, three of the tools are applied to the family as a whole (Browne & Saqi, 1988; Browne et al., 2006; Johnson, 1985) and two only apply to the mother (Brockington et al., 2001; Grietens et al., 2004). Only one tool is clear about when the assessment is carried out in relation to the child's age as it is based on four visits in the first year of life (Browne et al., 2006). Grietens et al. (2004) stated that the assessment is to be carried by the end of the designated visiting period (4

visits for the first child and 3 visits for those born after the first child). However, it was not made clear how old the child would be during this period.

Only one of the tools is clear about exactly how and where the assessment is completed (Browne et al., 2006). It is introduced during the first visit to a family as a way of involving the parents to explore and discuss their situations rather than used as a checklist to tick by health visitors or a questionnaire to fill in by the parents and is followed by observations and discussions about parental perceptions, attitudes and behaviour. This is in line with the principle of working in partnership with parents. The responses are not restricted to being from the mother or the father. It can be addressed to either or both of them, depending on which way the parents feel most comfortable or appropriate. In terms of training for assessors to ensure the correct and consistent use of the tools, only the CARE programme and Grietens et al. (2004) specified the specialist training sessions and materials provided to health visitors prior to carrying out any assessment.

Out of the five assessments, only the CARE programme (Browne et al., 2006) includes a structured health visiting intervention programme for families identified as 'high priority' for services. The intervention adopted by the CARE programme is the Parent Adviser Partnership (PAP) Programme, developed by Professor Hilton Davies independently of the development of the assessment component (Davies et al., 2002a; Davies et al., 2002b). This intervention programme has been widely adopted in the UK and evaluated in the Oxfordshire Intensive Home Visiting Study (Barlow et al., 2007; Barlow et al., 2003). Although the Oxfordshire Intensive Home Visiting Study is not to be concluded until the end of 2007, interim results have shown the programme to be effective (Barlow et al., 2007).

Brockington et al's (2001) assessment has also been implemented in a clinical setting. However, the project is not published and it is not possible to obtain the unpublished data. Other assessment systems have not gone beyond the initial stage of development.

Table 1: Characteristics of the validated assessment instruments

Authors	Brockington et al. (2001)	Browne&Saqi (1988), Browne (1995)	Browne et al. (2000, 2006)	Grietens et al. (2004)	Johnson (1985)
Name of instrument	None	Health Visitor Checklist	Index of Need, the CARE Programme	None	Langley Clinic health visiting project
Place of development	Birmingham, UK	Surrey, UK	Southend, Essex, UK	Belgium (Flemish)	Manchester, UK
Objective of assessment	Identifying mother-infant bonding disorders	Identification of families at high risk of child maltreatment	Identification of children and families in need	Identifying risk of physical child abuse and neglect	Identifying risk of child abuse
Designated assessor	Midwives or health visitors	Midwives and health visitors	Health visitors	Health visitors (social nurses)	Health visitors
Type of assessment	A 25-item questionnaire	A 12-item weighted checklist. The weighting was decided based on the prevalence of each factor.	Index of Need: a 14-item weighted index to assess level of needs. The weighting was decided based on the prevalence of each factor	A 20-item scale	A 20-item weighted checklist. The weighting was decided on the importance reported by previous literature.
Factors covered in the assessment	Scale 1: Impaired bonding Scale 2: Rejection and anger Scale 3: Confidence in ability to care for her baby Scale 4: Expressing the likelihood to hurt her baby or admitting to having caused harm	• Complications during birth/ separated from baby at birth (1) • Mother or partner under 21 years of age (1) • Mother or partner not biologically related to the child (1) • Twins or less than 18 months between births (1) • Child with physical or mental disabilities (1) • Parent abused as a child (2) • Serious financial problems (2) • Mother or partner treated for mental illness/ depression or alcohol/drug abuse (2) • Infant seriously ill, premature or weighed under 2.5kg at birth (2) • Single parent (3) • Adult in the household with violent tendencies (3) • Mother or partner feeling indifferent about their baby (3)	• Complications during birth/separated from baby at birth (1) • Mother or partner under 21 years of age (1) • Mother or partner not biologically related to the child (1) • Twins or less than 18 months between births (1) • Child with physical or mental disabilities (1) • Feelings of isolation (1) • Parent abused as a child (2) • Serious financial problems (2) • Mother or partner treated for mental illness or depression (2) • Dependency for drugs or alcohol (2) • Infant seriously ill, premature or weighed under 2.5kg at birth (2) • Single parent (3) • Adult in the household with violent tendencies (3) • Mother or partner feeling indifferent about their baby (3) In addition, an assessment is made over 4 home visits of • Parent's attitudes to child • Parent's perception of child • Quality of parenting • Development of infant attachment to parent These assessments contributed to the presence of absence of 'Mother or partner feeling indifferent about their baby' in the Index of Need.	Isolation • Mother intimates that she is alone facing the problem • Mother has few contacts outside the family and is dissatisfied with it • Mother is unable to adequately seek help or support • Mother is dissatisfied with contacts with family/friends • There is not much support from the partner • Mother has gloomy expectations • Mother intimates that she is unhappy • Pregnancy and delivery are very negatively spoken of Psychological complexity • Mother expects the baby to give abundant love • Mother speaks often about herself and not about the baby • Under stress, mother soon turns out to be helpless • Mother has already gone through several crises and it seems that she has difficulties getting over it • Mother does not show much self-confidence • Mother intimates that as a child, she did not get much love from her mother or family Communication problems • I have a feeling that the information mother is giving on how she deals with the baby is incomplete and does not tally. • Mother does not keep to the appointments regarding home visits and consultations • There is an atmosphere of secrecy in this family • I feel uncomfortable in this family • Mother does not take advice or only in part, on how to take care of the baby • Mother sets few limits and does not give much structure to baby's life	**Family background** • Social work involvement* • Evidence of abnormal relationships in extended family, e.g. abusive parents* • Socio-economic problems • Marital problems • Previous evidence of child abuse in the family • Male in the household not the father of the child* • 3 children in household under 5 years of age • Dissatisfaction or problems with the house **Mother** • mother under the age of 20 on birth of first child* • Emotional disturbances in the past • Capabilities of mother questioned by professionals in the past • Attitude to the baby during pregnancy, i.e. rejecting* • Mother educationally subnormal • Mother ill during first year of child's life* **Child** • Child admitted to S.C.B.U.* • Difficult delivery and pregnancy • Child ill in the first 6 months of life* • Difficult child e.g. overactive* • Difficult baby e.g. unsettled, colic, feeding difficulties etc. • Abnormal parental expectations of the child Each item with * receive a score of 2 when 1 score is given to the rest of the items.
Cut-off for 'in need' or 'high risk' category	Scale 1: score ≥ 12 Scale 2: score ≥ 17 Scale 3: score ≥ 10 Scale 4: score ≥ 3 The decision on the cut-off point was based on statistical analyses.	A total score ≥ 5 The decision on the cut-off point was based on statistical analyses.	A total score ≥ 5 The decision on the cut-off point was based on statistical analyses.	Not stated	A total score ≥ 8. The cut-off point was arbitrary, as it was NOT based on prior statistical analysis.

Table 2: The implementation of the validated assessment tools

Authors	Brockington et al. (2001)	Browne & Saqi (1988), Browne (1995)	Browne et al. (2006)	Grietens et al. (2004)	Johnson (1985)
Name of instrument	None	Health Visitor Checklist	Index of Need, the CARE Programme	None	None
Designated assessor	Midwives or health visitors	Midwives and health visitors	Health visitors	Health visitors (social nurses)	Health visitors
Instrument applied to	Mother	Mother	Family	Mother	Family
When the assessment is carried out	Not specified	At birth by the midwife and 10 days postnatally by the health visitor	Index of Need: 10 days after birth and at 12 months. Parenting observations at 10 days, 3, 6 and 12 months.	By the end of the 4 visits in case of the first child; By the end of the 3 visits for those born after the first child.	Not specified
Venue of assessment	Not specified	Hospital and family home	Family home	For social nurses to complete outside the client contact time	Not specified
How the assessment was carried out	Not specified	Midwife identified which factors were present of the 12 items listed . The checklist was passed over to the health visitor 10 days after birth and health visitor (not blind to midwife's ratings) rated which factors were presented again.	Index of Need is NOT supposed to be used as a checklist to tick or a list of questions to ask directly. It is introduced during the first visits to the family as a way of involving the parents to explore and discuss their situations. They may delay discussing it until they feel safe and confident in their relationship with HV or other professionals. Index of Need is not designed for the identification of children likely to suffer significant harm in isolation of the parenting observation.	Social nurses respond to all items in the scale based on any available information and observation during visits	Not specified
Intervention for families categorised as 'in need' or 'high risk'	Not specified	Increased support from health visitors but no structured intervention programme	High priority for services and Parent Adviser Partnership Programme	No	42% of 'at risk' families (n=130) received more input from health visitors but no designated or structured intervention was mentioned or proposed as future arrangement.
Training course to ensure use of the tool	Not stated	Not stated	A 3 day workshop on the assessment package, inc. 1 day on Index of Need and 2 days on behavioural observation led by qualified psychologists, accompanied by training packs. Further training is provided on the intervention and joint referral criteria with social services and other professional agencies.	Training was provided on • Understanding of parenting and child maltreatment and • The use and scoring of the initial version in order for them to carry out research However, it is not specified whether there is formal training for the validated version.	Not stated
Training materials	Not stated	Not stated	2 training packs (Hegarty, 2000a; Hegarty, 2000b) given out during the training. 1 Book entitled giving theoretical background, detailed overview of the programme and research evidence (Browne et al., 2006)	Training pack for the use and scoring of the initial non-standardised, non-validated version.	Not stated
Validation & evaluation design	Cross-sectional study: Comparing scores on all scales against results gained from a 'reference standard' test.	A prospective cohort study	A prospective cohort study Assessment applied to a cohort and follow up on prospectively.	Case control study: Investigating scores and differences between abusers and non-abusers	Possibly a cohort study, as the time point of assessment was not specified. The time gap between the assessment and outcome was not specified.

Non-validated assessment tools

Of the three non-validated instruments, Barlow et al. (2003) and the Health Needs Assessment Tools (HNAT) were used in the UK and the other was used in Italy (Ammaniti et al., 2006). HNAT is a standard assessment procedure carried out by health visitors under guidelines from the National Health Service (2001) but there is no published evidence for its validation and evaluation. Cowley & Huston (2003) conducted a qualitative study to gather views on HNAT from 30 health visitors and 19 mothers and analysed 21 video taped assessment sessions to see how mothers actually respond in the assessment. They concluded that the structured format encouraged health visitors to question rather than listen to their clients and the assessment caused anxiety and distress to vulnerable clients, particularly the most vulnerable ones.

The assessment tools by Ammaniti et al. (2006) and Barlow et al. (2003) were both 'entry criteria' for the randomised controlled trials they conducted for the evaluation of a health visiting intervention programme and were not standard practice adopted by the authorities. Ammaniti et al. (2006) did not specify who carried out the assessment whereas Barlow et al. (2003) used midwives for recruitment and assessment although social workers were also approached for referrals when they could not get enough participants through midwives. Barlow et al. (2004) explained that the reason why they did not validate and evaluate the assessment tool was that they would not have access to follow up on all the women who did not meet the entry criteria as they were excluded from their studies. Without a comparison group, assessment results and later outcomes were impossible to evaluate.

All three tools are short in length, with 9 factors covered in Ammaniti et al. (2006), 9 factors assessed in Barlow et al. (2003) and 18 items in the HNAT. All of them look into psychosocial, socioeconomic and health factors. Both Ammaniti et al. (2006) and Barlow et al. (2003) clearly stated the threshold for a participant to be entered into the trial whereas HNAT did not have a set cut off point, as health visitors are expected to discuss and make a decision on the level of needs together with the parent.

In terms of the actual assessment, all three tools are administered to the mother. Ammaniti et al. (2006) and Barlow et al. (2003) both had the assessment carried out during the second trimester of pregnancy. HNAT did not dictate when the assessment takes place. Rather, it is up to the health visitors to decide when the best time is to introduce the assessment. Both Barlow et al. (2003) and Cowley & Huston (2003) stated that their assessments take place during client contact. However, it was not clear exactly where it actually happens. Normally, midwives do not do home visits unless there are obvious concerns about the pregnancy and women are expected to attend routine check ups in clinics. Health visitors in the UK are supposed to make home visits after birth. However in some parts of the UK, health visiting services have become more clinic based and the frequency of visits have reduced to only once. Therefore, it cannot be assume that HNAT is administered in the family home during visits.

It was stated in Cowley & Huston (2003) that section A of the HNAT is filled in by the mother, second B is completed by the mother together with the health visitor and section C is filled out by the mother but as mentioned before, the final decision on the level of need (high, medium or low) is made jointly with the parent. This is in line with the 'working in partnership' principle. Similarly, Barlow et al. (2003) asked midwives to carry out assessment during routine check up and consultation and gain consent for the women to enter their trial if they recommended the mother to 'receive more help'. Those who are approached could decline without any consequences on the standard services they are entitled to. What is not clear is how the information from the mother is extracted, i.e. whether by a structured interview and/or information from. Ammaniti et al. (2006) reported that interviews were carried out to gather information but no details were given as to whether it was structured.

All three systems responded to clients' need determined by the assessment in some way. Ammaniti et al. (2006) and Barlow et al. (2003) used RCTs to evaluate structured intervention programmes. Therefore, the interventions were the main focus of their studies and naturally serve as the response to a high level of need. By contrast, the HNAT had no structured intervention for families with high needs. The assessment results simply helped health visitors adjust the level of support they offered to families. Finally, none of the programmes specify whether training sessions and materials are provided to assessors to ensure the correct understanding and consistent use of the tools.

Table 3: Characteristics of the non-validated assessment instruments

Authors	Ammaniti et al., (2006)	Barlow et al. (2003)	Cowley & Houston (2003)
Name of instrument	None	None	**Health Needs Assessment Tools (HNAT)**
Country of development	Italy	UK	UK
Objective of assessment	Identifying risk of poor mother-infant relationship	Identifying risk of child maltreatment	Identifying health needs of child
Designated assessor	Not specified	Midwives	Health visitors
Type of assessment	9 psychosocial risk factors & depression (unweighted)	9 psychosocial risk factors commonly associated with child maltreatment (unweighted)	18 items which cover several major risk factors

Factors covered in the assessment	**Interview:**
	• Low educational level
	• Low socioeconomic status
	• Single parenthood
	• Family psychiatric history
	• History of physical or sexual abuse
	• Antisocial behaviour
	• Stressful life events (e.g. loss, separation, abortion)
	• Large family
	• Lack of social support
	
	Psychometric assessment:
	• Depression, measured by Center for Epidemiological Studies-Depression Scale (CES-D, Radloff, 1977)
		• Homelessness or moved home more than twice in the last 12 months
		• Severe debt/financial hardship
		• Absence of support networks
		• Aged 17 years or under
		• Serious mental illness requiring treatment (current or previous)
		• Moderate learning disabilities
		• Substance misuse/addiction (current or previous)
		• Domestic violence
		• Social work involvement relating to children (current or previous)
		• Non-specific serious concern expressed by midwife
			Section A: Index of Need
			Health
			• Either parent has a serious illness or concerns about a hereditary condition in their family
			• Child was a) born premature, b) has a serious illness or c) was separated from mother at birth
			• Either parent has a child with special needs
			• Either parent has had mental health/psychological problems or depression
			• Either parent feel they or their partner have a dependency on alcohol or drugs
			
			Environment
			• Parent has concern about housing or environment
			• Parent dissatisfied with level of support from family and friends
			
			Social
			• Parent has experienced a major change in family situation such as
			  - Bereavement
			  - Acquired refugee status or an asylum seeker
			  - Separation from a partner
			  - Frequent changes of address
			  - Homelessness
			  - Other
			• Parent did not think it was a good time to become a parent
			• Either parent had difficult experiences as a child which affects their parenting
			• There is an adult in the house with violent tendencies
			• Either parent has difficulties

Authors	Ammaniti et al., (2006)	Barlow et al. (2003)	Cowley & Houston (2003)
			in understanding English • Either parent has serious financial worries **Section B: Assessment interview with parent/carer completed with parent/carer** 1. how loving the parent feels towards the child 2. how happy the parent is with the child's behaviour 3. how confident the parent feels about being a parent 4. how happy the parent is with the child's progress **Section C: Intuitive assessment completed by HVs** • factors that lead HV to think this family may need more or less input
Cut-off for 'in need' or 'high risk' category	a) Depression (a score of 20* or above) on CES-D AND 0-1 psychosocial risk factor OR b) Low depressive symptoms but with at least 3 psychosocial risk factor	The presence of any of the above risk factors	No cut-off point is given. HV and parents decide together which level of need the family has (high, medium or low)

Table 4: The implementation of the non-validated studies

	Ammaniti et al., (2006)	Barlow et al. (2003)	Cowley & Houston (2003)
Name of instrument			**Health Needs Assessment Tools**
Country of development	Italy	UK	UK
Designated assessor	Not specified	Midwives	Health visitors
Instrument administered to	Mother	Mother	Mother
When the assessment is carried out	Second trimester of pregnancy	Second trimester of pregnancy	HV decides when to introduce the assessment
Venue of assessment	Not stated	During client contact	During client contact
How the assessment is carried out	Interviews were carried out to gather information. No further details were given on whether it was structured.	No details were given on how exactly the information was extracted.	The checklist is handed to the parents for them to complete where they can and HV work together with parents on section B.
Intervention for families categorised as 'in need' or 'high risk'	Home Visiting Program	Parent Adviser Partnership Programme	Increased input from health visitors, but no structured programme is specified.
Training course to ensure use of the tool	Not stated	Not stated	Not stated
Training materials	Not stated	Not stated	Not stated
Validation & evaluation design	None	None	Cowley & Huston (2003) carried out a qualitative study to gather views on the instrument from a group of 30 visitors and 19 mothers and observe 21 video taped assessment sessions.

Methods of evaluation for the validated assessments

To be able to make valid comparison between assessment tools, it is important to look into the validation and evaluation studies for those instruments. Naturally, this does not apply to instruments that have not been validated. There are five studies, reported in seven publications, for the four instruments (Brockington et al., 2001; Browne, 1995; Browne et al., 2006; Browne, Hamilton, Hegarty, & Blissett, 2000; Browne et al., 1988; Grietens et al., 2004; Johnson, 1985).

Of the five studies, only two (Browne & Saqi, 1988; Browne, 1995; Browne et al., 2000, 2006) were definitely cohort studies where participants were assessed first and then followed up for their outcomes. Johnson (1985) may be a cohort study or it could simply be a cross sectional where the assessment and the outcomes were measured at the same time. The time points of when each measure was taken were not stated and the reporting was not explicit enough to decide. Table 5 summarises the characteristics of the cohort studies, detailing the validation and evaluation of the two instruments.

Brockington et al. (2001) and Grietens et al., (2004) on the other hand, both adopted case control designs where a group of individuals with the outcome and the other without were identified and the assessments were applied to the two different groups in order to compare the differences in the results between the groups. This approach is usually used in pilot studies to develop and validate the assessments. Both instruments were only finalised and reached their current forms after these studies but no published data are available to suggest a cohort design has been conducted to confirm their predictive validity. In the authors' knowledge, Brockington et al. (2001) has piloted their assessment scale with nurses but the study was not published and therefore not available for analysis. Table 6 summarises the characteristics the case control studies detailing the development and validation of the two validated instruments.

It is not possible to give the exact total number of the participants as some studies counted the total number of births without specifying how many of them were twins (Browne & Saqi, 1988; Browne, 1995; Browne et al., 2000, 2006) and some reported the number of families visited (Grietens et al., 2004; Johnson, 1985). Only Brockington et al. (2002) made both clear. Nevertheless, the figures reported in each study can still reflect the size of each study without making the distinctions between the number of families and the number of births. It ranges from 104 births in Brockington et al. (2002) to 14,252 births (Browne & Saqi; Browne, 1995).

Table 5: Cohort study methods for evaluation of assessment procedures

Authors	Browne & Saqi (1988) Browne (1995)	Browne et al. (2000, 2006)	Johnson (1985)
Name of instrument	Health Visitor Checklist	Index of Need	Langley Clinic health visiting project
Location of study	Surrey Primary Health Authority, UK	Southend on sea, under Essex Primary Health Authority, UK	Rochdale, Manchester, UK
Period of recruitment & Period of entire data collection	1 Jan 1985 – 31 Dec 1986 (2 years) 1 Jan 1985 – 31 Dec 1991	1 Apr 1995 – 30 Jun 1998 (3 years) 1 Apr 1995 – 31 Jul 1999	Not stated Outcome was taken in Jan 1984
Final sample size	14,252 births (the entire birth cohort except 5% who declined or dropped out)	4,351 births (the entire birth cohort except 9% who declined or dropped out)	1,064 families
Approached but did not complete study	747 declined	310 declined 114 left during the first sixth months	Not stated
Sample characteristics			
Age Ethnicity SES Other	Not stated Not stated Above average of the country Either parent reported a history of physical and/or sexual maltreatment in their own childhood in 7.6% (1,083) of the families.	Not stated 95.1% white UK child 1.4% white European child 1.3% mixed ethnicity 1.2% Asian child 0.5% Afro-Caribbean 0.02% South American* Below the average of the country Either parent reported a history of physical and/or sexual maltreatment in their own childhood in 3% (135) of the families.* 25 maltreated children under 13 months 11 neglected 9 physically abused 4 emotionally abused 1 sexually abused	Not stated Not stated The area where the study was carried out was a socio-economically deprived estate in a major city. However, no details on this particular sample was given. Not stated
Outcome measures	Attendance of a case conference for suspected or actual child maltreatment	Referrals to social services for suspected or actual maltreatment	Suspected or actual maltreatment observed by health visitors Referrals to child protection agencies
Time of initial assessment and by whom	Midwives assesses family at birth and health visitors adds any information during the first months of child's life	Around time of birth, at the new birth visit	Postnatal, exact time point not specified
Age of child at assessment of outcome	5 years	13 months	Not stated
Length of follow up	5 years	13 months	Not stated
Intervention for families categorised as 'high risk' (HR)	Increased support from health visitors	High priority for services and the Parent Adviser Partnership Programme	No
Proportion of sample completed the study	95%	91%	Not stated
Were those who did not complete the study similar to those who did?	Not stated	Yes	Not stated
* calculated from information provided in (Dixon, Browne, & Hamilton-Giachritsis, 2005)			

Table 6:   Case control study methods for evaluation of assessment procedures

Authors	Brockington et al. (2002)	Grietens et al. (2004)
Name of in-strument	None	None
Location of study	Birmingham pregnancy-related sub-regional service (clinic), UK	Area not named or described, Belgium (Flemish)
Period of re-cruitment & Period of entire data collection	Not stated Not stated	3 months but no exact dates were given Oct 1999 – Mar 2000
Final sample size	*101 mothers with 104 babies as 3 had twins	391 parents
Approached but did not complete study	Not stated	901 parents for whom the social nurses had not completed all standard visits within the 3 month study registration period
Sample charac-teristics		***Abusive/neglectful mothers (n=18) vs. Non abusive (n=373)***
Age	Not stated	Range 19-41 vs. 22-44, Mean 29.46 vs. 29.77, SD 1.29 vs. 3.50
Ethnicity	Not stated	81.8% Belgium, the rest are of European, African or Asian origins vs. 77.8% Belgium, the rest are of European origin
SES	Not stated	
Other	33 normal mothers recruited from GPs or obstetric clinic 22 mothers of high risk pregnancies 21 depressed mothers with a normal mother infant relationship 28 depressed mothers with impaired bond-ing	The abusive group scored significantly higher on 'social deprivation scale' than the nont-abusive (p<.01). The duration of pregnancy in the abusive group was significantly shorter than non-abusive (p<.01). ***birth order*** Range 1-4 vs. 1-7 Mean 2.17 vs. 1.94 SD 1.42 vs. 0.42
Outcome measures	'Reference standard' test: a 24-probe section on mother-infant relationship in the 'Birmingham Interview for Maternal Mental Health'	Diagnosed by Confidential Doctors' Teams (multidisciplinary)
Time of initial assessment and by whom	Postnatal, personnel of assessment and exactly point of time not specified	By the end of 4 standard nurse visits for the first child By the end of 3 standard nurse visits
Age of child at assessment of outcome	Not stated	Not stated
Intervention for families cate-gorised as 'high risk' (HR)	Clinic based intervention but the type of treatment was not specified.	No
Proportion of sample com-pleted the study	N.A.	43.3%
Were those who did not complete the study similar to those who did?	Not stated	Not stated

Cohort studies (Table 5)

The size of the sample is affected by the recruitment method and period. All cohort studies (Browne & Saqi, 1988; Browne, 1995; Browne et al., 2000,

2006; Johnson, 1985) have large sample sizes, as cohorts of all the live births in an area were included. The first 'Health Visitor checklist' study (Browne & Saqi, 1988; Browne, 1995) was carried out with 14,252 births in Surrey, an area in the south of England normally distributed for social class. Ninety-five percent of families with a live birth between 1 January 1985 and 31 December 1986 in the area were assessed at birth and then followed up five years. There is no economic and health profile during the time of the study readily available. However, according to the health profiles in 2006 (APHO & Department of Health, 2006), the overall poverty rate in surrey is low, the educational attainment level is significantly high and the rate of infant deaths are also significantly lower than the average for England.

The second study (Browne et al., 2000, 2006) was carried out with 4,351 births in a Southend on Sea, Essex, a working class seaside resort in the south east of England. Ninety-one percent of families with a live birth between 1 April 1995 and 30 June 1998 were included and the follow up period was 13 months after an initial postnatal assessment. As reported in the health profiles in 2007 (APHO & Department of Health, 2007c), one in seven adults claim benefits and one in four children live in households relying on benefits. The rate of infant death is the national average but there is a higher rate of teenage pregnancy.

The study by Johnson (1985) was conducted in a deprived area in Manchester, England. No descriptions of the particular area at the time of study was provided but the current health profile (APHO & Department of Health, 2007b) states that Manchester has higher rates of infant deaths and teenage pregnancy than the national average. All the health visitors at a clinic participated and all the live births in their caseloads were investigated during the study period which ended in January 1984 (N=1,064). This means that all the families came from the same area. However, it is unclear whether all the participating health visitors covered all the families or just part of families in the area and when they carried out the assessment. It is also unclear which birth cohort was taken and the length of follow up or time between the assessment and the measure of the outcome. She also did not report the number of attrition after recruitment.

In terms of sample demographics and characteristics, two out of the three studies did not report any information (Browne & Saqi, 1988; Johnson, 1985). Only one study (Browne et al., 2006) reported ethnicity and types of maltreatment. Only the CARE programme (Browne et al., 2006) confirmed that those who did not complete the study similar or different to those who did.

When determining the outcome of participants, Johnson (1985) included suspected or actual maltreatment both observed by health visitors and actual referrals to child protection agencies whereas the other two cohort studies (Browne & Saqi, 1988; Browne et al., 2006) adapted stricter measures of using actual attendance of a case conference for suspected or actual child maltreatment and referrals to social services for suspected or actual maltreatment (Browne et al., 2006).

Out of the three cohort studies, only the second study of the CARE programme (Browne et al., 2006) had a structured intervention designated for

families classified as in high need. The earlier studies (Browne & Saqi, 1988; Johnson, 1985) did not have such provision.

Case control studies (Table 6)

Brockington et al. (2001), a case control study, is based on a clinical sample (specialist services), in Birmingham, England. The area, similar to Manchester with an urban population of three million, has high rates of infant deaths and teenage pregnancy (APHO & Department of Health, 2007a). A small sample of 101 mothers were assessed but those initially approached and the number that did not consent or left the study were not reported. The duration of data collection was not made clear in their publication. Therefore, it is not possible to decide whether a longer duration of recruitment and data collection would have increased the sample size.

Grietens et al. (2004) also had a relatively small sample size of 391 mothers although they attempted to recruit all mothers with a newborn who were visited by 40 social nurses in an unspecified area in Belgium. The recruitment period of the study was only 3 months and it was impossible to recruit all 901 in that time. Hence, it needs to make clear unclear whether the sample is representative of the overall population. For example, it is unclear whether the 40 participating nurses sampled the entire area or whether they recruited from the same area.

In terms of the sample demographics, Brockington et al. (2001) did not provide any information. Grietens et al. (2004) reported that the age and ethnicity of groups and found greater social deprivation in the abuse/neglect group.

With regard to sample characteristics, Brockington et al. (2001) reported that 32.6% of the mothers were normal, 22.1% had high risk pregnancies, 21.7% were depressed with normal bonding and 27.7% were depressed with impaired bonding. Grietens et al. (2004), on the other hand, paid attention to the duration of pregnancy and birth orders. It was found that the duration of pregnancy was significantly shorter in the abusive group compared to the non-abusive group and the mean birth order.

When determining the outcomes, Brockington et al. (2001) used a reference test, Birmingham Interview for Maternal Mental Health, to decide whether a mother truly had impaired bonding. Diagnoses were decided by two psychiatrists and any differences in the opinion were resolved by discussion. Grietens et al. (2004) used the decisions made by the multidisciplinary 'Confidential Doctors' Teams' on whether parents are abusive or non-abusive. Neither study specified the exact time of assessment nor the similarity or differences between those who completed the studies and those who did not. Neither study stated any intervention that would be provided to families identified as high risk by their tools.

Quality of the included studies

The findings of the above studies should be interpreted in relation to the quality of the research that produces the evidence. The cohort studies are considered stronger evidence partly because of the prospective nature of the outcome, so at the time of the assessment, the nurse is blind as to the outcome for the families. However, in later referrals to child protection agencies by health visitors, the nurse was not blind to the high or low risk status of the family.

Cohort studies (Table 7)

Before looking into the studies, a fundamental question needs to be asked about how the items in the instruments were selected. Both the 'Health Visitor checklist' and the 'Index of Need' consist of risk factors commonly found to be associated with child maltreatment in previous research. The inclusion and weighing of those items were based on the prevalence of each item in the population to which it applies to (Browne & Saqi, 1988; Browne et al., 2006). On the other hand, Johnson (1985) included risk factors commonly found to be associated with child maltreatment in previous literature. However, no effort was made to establish the prevalence of each item in the targeted population and the weighing of those items were based on the importance stipulated by previous research in other parts of the world.

All the cohort studies have a clear description of the instruments and cut-off points. They are also clear about the methods used to identify whether maltreatment had occurred. The 'Health Visitor checklist' study used attendances of case conferences due to actual or suspected child maltreatment (Browne & Saqi, 1988; Browne, 1995) and the CARE programme also included referrals to social services as an indicator (Browne et al., 2000; 2006). However, Johnson (1985) relied on either maltreatment determined by the participating health visitors or actual referrals to social services, NSPCC or other agencies. Data from Johnson (1985) shows that the total number of families seen as maltreating was 91 but the total number of cases referred to child protection agencies was only 59. This means the number of actual referrals to child protection agencies may be an underestimate of the total number of maltreating families, especially when studies only use social services referrals.

All three studies had a seemingly large sample size. However, as the base rate of child maltreatment is very small it is necessary for a cohort study to have a large sample so that the number of families with the outcome of interest would be sufficient to make the statistical analysis meaningful. Therefore, the sample size of Johnson (1985) may be too small. Similarly, the follow up period is too short with the exception of that reported in the 'Health Visitor checklist' study (Browne and Saqi, 1988; Browne, 1995).

In terms of reporting, Johnson (1985) did not consider the implications of sensitivity and specificity, which are considered important indicators to the test performance. Nevertheless, the raw figures were presented in the tables in the paper so that the readers can calculate the relevant statistics

and compare them to those sensitivity and specificity reported in the other cohort studies. Table 7 summarises the overall quality of study and reporting across the cohort studies. It ranges from poor (Johnson, 1985) to reasonable (Browne & Saqi, 1988; Browne, 1995; Browne et al., 2000; 2006). Out of the three, Browne et al. (2000, 2006) is considered the best, as it employed more measures than the others to reduce bias. However, the follow up period could have been longer to increase the applicability of the instrument and more background information should be given in order to assess the generalisability of the instrument. Follow up periods of one or two years are defended by the fact that the majority of fatal physically abuse and neglect occurs during infancy (Hamilton & Browne, 2002).

Table 7: The quality of study and reporting in the cohort studies

	Browne & Saqi (1988); Browne & Herbert (1997)	Browne et al. (2000; 2006)	Johnson (1985)
Risk factors selected by acceptable methods	Yes	Yes	No
Adequate sample size	Yes	Yes	Partially yes
Sufficient sample descriptions	No	Partially yes	No
Description of the tool detailed enough	Yes	Yes	Partially yes
Outcome measures clearly defined	Yes	Yes	Yes
Cut-off points decided by acceptable methods	Yes	Yes	No
Measures to ensure consistent use by health visitors	Yes	Yes	Unclear
Blinding of outcome assessors	Partially yes*	Partially yes*	No
Adequate follow up period	Yes	Partially yes	Unclear
Attrition dealt with	No	Partially yes	No
Appropriate statistical analysis	Yes	Yes	Partially yes
Overall quality	Reasonable	Reasonable	Poor

* for the child protection referrals not made by the health visitor

Case control studies (Table 8)

Brockington et al. (2001) attempted to merge two existing scales into one. However, it was not clear how the two existing scales were formulated in the first place. Grietens et al. (2004) on the other hand adopted the same approach as Browne and Saqi (1988) and Browne et al. (2000, 2006) when putting their initial questionnaire together. They also used focus group to

gather expert opinions on what should be included and how items should be worded. Nevertheless, both studies were clear on the measures to determine the outcomes investigated in their studies.

Brockington et al. (2001) simply reported the total number of participants but did not state the total number approached initially or the total number who did not give consent to participate. There was a lack of background information on the sample and no details were given about any differences between those who completed the study and those who did not. It should be noted that the majority of the mothers in the sample were depressed (50%) in their relatively small sample, which is higher than the prevalence of the condition in a general population. Therefore, the current findings can only be applied to a clinical population. On the other hand, Grietens et al. (2004) had a relatively large sample size. The total number recruited within the period of their study was reported against the total caseload to give an idea of the questionable representativeness of the sample. However, attrition was not addressed in their study either.

Similarly, neither study specified whether the outcome assessors were blind to the results from and actual responses to the scales. In Brockington et al. (2001), two of the authors assessed the outcome of those mothers and it was not clear whether they were kept blind from all the questionnaire results. In Grietens et al. (2004), the 'diagnoses' of child maltreatment was made by the multidisciplinary 'Confidential Doctors' Teams'. It is not clear how those teams work, whether any member of those teams could have had access to the results from the questionnaires.

However, Grietens et al. (2004) should be commended for the provision of training for the participating social nurses to ensure correct and consistent administration of the questionnaire whereas Brockington et al. (2001) simply did not provide any information related to training. It would be more helpful if the cut-off point for the scale in Grietens et al. (2004) was explored and the performance of the scale based on the cut-off point was presented. Overall, Grietens et al. (2004) is considered to have reached reasonable quality of study and reporting whereas Brockington et al. (2001) is considered of poor quality.

Table 8: The quality of study and reporting in the cohort studies

	Brockington et al. (2001)	Grietens et al. (2004)
Risk factors selected by acceptable methods	Unclear	Yes
Adequate sample size	No	Yes
Sufficient sample descriptions	Partially yes	Yes
Description of the tool detailed enough	Partially yes	Partially yes
Outcome measures clearly defined	Yes	Yes
Cut-off points decided by acceptable methods	Unclear	No
Measures to ensure consistent use by health visitors	Unclear	Yes
Blinding of outcome assessors	Unclear	Unclear
Adequate follow up period	Not applicable	Not applicable
Attrition dealt with	Unclear	Unclear
Appropriate statistical analysis	Partially yes	Partially yes
Overall quality	Poor	Reasonable

Overall performance of the assessment procedures

Cohort studies (Table 9)

In the UK cohort studies, the incidence of child maltreatment ranges from 0.6% to 0.9% and the proportion of sample classified as high need or high risk ranges from 4% to 7%. Johnson (1985) reported a slightly higher incidence of 0.9% than Browne & Saqi (1988) who found 0.7% and Browne et al. (2006) who found 0.6%. This could be due to a number of reasons such as the differences between outcome measures they adopted and variations in the referral criteria. This in turn affects the sensitivity and specificity.

Browne and Saqi (1988) reported 68% sensitivity and 94% specificity for their 12 item checklist. Similarly, Browne et al. (2006) reported 70% sensitivity and 96% specificity for their 14 item checklist, with the additional observations of parenting objectively influencing the health visitors rating of the 'parent insensitive' risk factor on the Index of Need. However, while these assessments demonstrate reasonable ability to correctly identified families in high or low need, the high false positive rates mean the PPVs are both extremely low at 7% and 11% respectively. This means a high score on the Index cannot indicate probability of high risk in reality. When applied to

a large population, the actual number of false positives would be the vast majority of the high risk group (93% and 89% respectively). Johnson (1985) seemed to have performed better as it achieved 89% sensitivity, 95% specificity and the PPV was 63% and the false positives were in the majority of the high risk group (47%). However, unlike Browne and Herbert (1997) and Browne et al. (2006), the follow up period was not specified for Johnson (1985). Therefore, it is not clear how long the predictive value applies to. Also due to the wider ranges of measures taken for the outcome, the results between the Johnson (1985) and the other studies are not entirely comparable. Furthermore, the demographics of the areas under investigation may play a role in the differences in the results. It seems that the more deprived the area, the greater the PPV. Another factor that could make a difference was the way the assessment was carried out. The Index of Need is completed in partnership with the parents and health visitors in the CARE programme were specially trained to use the tool and engage their clients but no such information was reported in Johnson (1985). The results of all the cohort studies are presented in table 9.

Case control studies

The two studies included in this review are of different focuses (see Table 8 for the findings of these two studies). Brockington et al. (2001) was designed to identify bonding disorders between the mother and the infant whereas Grietens et al. (2004) developed a scale to identify parents at risk of maltreating their children.

Brockington et al. (2001) reported that the scale 1 of the measure was effective in identifying mild bonding disorders as it reached 93% specificity and 84% specificity and the PPV was 74% and the NPV was 96%. The scale was also sensitive to those with severe bonding disorder, as it reached 100% sensitivity, 84% specificity; the PPV was 67% and the NPV was 100%. However, there are some problems with the analyses. First of all, they did not analyse the results with the entire sample. They considered normal mothers and all the mothers diagnosed as having any mild form of bonding disorders by the reference test (i.e. excluding those with severe bonding disorders) separately to the analysis for the accuracy of the scales for identifying mild bonding disorders. Then, they entered normal mothers and those diagnosed with severe bonding disorders (i.e. excluding those with mild disorders) into the analysis for severe bonding disorders. This means the figures only show the scale's ability to identify mild/severe bonding disorders in the absence of one another. Therefore, the practical utility and ability of this scale to identify a mother with bonding disorders out of a population is highly questionable. A more fundamental problem was that the sample was purposely selected from clinics for the study and the majority of the sample suffered from depression. The incidence of mild bonding disorders in their sample was reported to be 34% in the sample for mild bonding disorders and 25 in the sample made up for severe bonding disorders, which are both higher than the prevalence in the general population. As the sample was not representative and consisted primarily of a clinical population, the findings cannot be generalised to normal population. Furthermore, it is un-

clear how the cut-off points for their scales were decided. The overall quality of study and reporting in Brockinton et al. (2001), therefore, is considered poor.

Grietens et al. (2004) conducted logical regression in their study and reported that high scores on communication problems and isolation were predictive of high scores on social deprivation, which is highly associated with child maltreatment. The quality of study of reporting in Grietens et al. (2004) was reasonable and the scale they developed seem worthy of future follow up to ascertain the feasibility of implementation in the real world. However, the scale seems to be reliant on health visitors' judgement and some of the items appeared to be quite abstract and subjective, for example, 'There is an atmosphere of secrecy in this family, 'I feel uncomfortable in this family'. Further details are needed to determine how health visitors are supposed to rate each item and what cues or evidence they need to see to make a decision, especially on the more subjective and abstract factors.

Table 9: Performance evaluation from the cohort studies

Study & contingency table				% of total sample assessed	Incidence of disorder or referrals	% of sample in high risk/ need group	Sensitivity	Specificity	PPV	NPV	Quality of study and reporting
Browne & Saqi (1988), Browne (1995)				100%	0.7%	7%	68%	94%	7%	99%	Reasonable
	Mal-treated	Non mal-treated	Total								
H risk	72	892	964								
L risk	34	13,254	13,288								
Total	106	14,146	14,252								
Browne et al (2000; 2006)				100%	0.6%	4%	70%	96%	11%	99%	Reasonable
	Re-ferred	Not referred	Total								
H need	19	157	176								
L need	8	4,167	4,175								
Total	27	4,324	4,351								
Johnson (1985) by HV classification				100%	0.9%	12%	89%	95%	63%	99%	Poor
	mal-treated	Not mal-treated	Total								
H risk	82	48	130								
L risk	9	925	934								
Total	91	973	1,064								

Conclusions and recommendations

A total number of eight assessment systems were reviewed. The overall evaluation of the non-validated tools are summarised in table 10 and that of the validated tools are illustrated in table 11.

Due to the absence of validation studies behind them, the three non-validated assessments by Ammaniti et al. (2006), Barlow et al. (2003) and HNAT cannot be properly evaluated. Both Ammaniti et al. (2006) and Barlow et al. (2003) were primarily evaluating intervention programmes and the assessment component was not the focus. The qualitative study by Cowley and Houston (2003) also raised issues on the use of the tool. The negative user feedback may be more related to the format of the assessment, the way it is carried out and the need for more training provided to health visitors. Therefore, it is not recommended for the three systems to be considered for practical use without further research.

All of the five structured assessments are multifactorial, as child maltreatment is a complex problem involving numerous factors and facets. The assessments are all of a reasonable length to make them feasible and practical and all reported a certain level of accuracy. However, it is important to note that none of the assessment instruments reported over 90% sensitivity except Brockington et al. (2001) whose primary focus was actually maternal bonding on a small clinical sample. Therefore, it is inappropriate to use any of the assessments to assess the 'risk' of child maltreatment. Nevertheless, with a reasonable level of accuracy, they can be used as an assessment on need in families. Prioritising families for more intensive support may in turn prevent child maltreatment as the 'sensitivity' shows that the majority of maltreating parents are in the high need (or priority) group. Grietens et al. (2004) did not analyse the sensitivity and specificity of their finalised scale. The current validity to another scale that distinguishes maltreating mothers to non-maltreating ones is not enough to justify its accuracy. A prospective cohort study is required to ascertain the accuracy and usefulness of this tool.

The CARE programme (Browne et al, 2000; 2006) with its 14 item Index of Need, parenting observations and Parent Adviser Partnership programme (Davies et al., 2002a; Davies et al., 2002b) shows promise. The formation of this programme came about differently from Olds et al., (1986) and Barlow et al. (2003) where the main aim was to evaluate the effectiveness of intervention.

The 12 item 'Health Visitor checklist' (Browne & Saqi, 1988) and the related are CARE programme (Browne et al., 2000; 2006) is the only assessment systems that have undergone long term follow up (5 years and 13 months). Not only has it shown reasonable accuracy, health visitors have also felt positive about it, as the way it is carried out makes sure that it serves as a catalyst encourage parents to talk about family difficulties rather than a hindrance in the process of prioritising families in need (Hamilton et al., 2002). More importantly, the assessment results help health visitors prioritise cases on an object basis better and enable to them take actions early and make appropriate referrals to other agencies (e.g. mental health professionals, social services) for prevention work.

The early Health Visitor Checklist (Browne & Saqi, 1988; Browne, 1995) was criticised as a 'tick list' of questions to read out and therefore like many other checklists was considered as 'too mechanical' and an obstacle to engage parents (Appleton, 1994). However, most health visitors actually welcome some form of assessment to help them prioritise cases when working with the limited resources. The 14 item Index of Need has been used in the CARE programme as a medium to engage parents. Trained health visitors introduced the Index of Need during visits as factors to discuss and explore if the parents wish to and whenever the parents feel ready. Some parents actually welcome this as an opportunity to talk about problems in the family previously not mentioned (Browne et al., 2006). The principle of the CARE programme, like the successful NFP programme in the USA, is to work in partnership with parents and their families.

Table 10: An overall evaluation of non-validated tools

Authors	Ammaniti et al. (2006)	Barlow et al. (2003)	NHS, UK
Name of instrument	None	None	HNAT
Strong theoretical grounding for the formation of the tool	Unclear	Yes	Unclear
Empirical evidence for the formation of the tool	Unclear	Unclear	Unclear
Sufficient details of the assessment			
The actual instrument	No	Yes	Yes
Time of assessment	Yes	Yes	Yes
Who are assessed	Yes	Yes	Yes
The way the assessment is carried out	Partially yes	No	unclear
Reasonable length	Yes	Yes	Yes
Training	Unclear	Unclear	Unclear
Cut-off clearly specified	Yes	Yes	Yes
Practical evaluation	No	No	No

Table 11: An overall evaluation of validated tools

Authors	Brocking-ton et al. (2001)	Browne & Saqi (1988), Browne (1995)	Browne et al. (2000; 2006)	Grietens et al. (2004)	Johnson (1985)
Name of instru-ment	None	Health Visitor Checklist	Index of Need	None	Langley Clinic health visiting project
Strong theoretical grounding for the formation of the tool	Unclear	Yes	Yes	Yes	Yes
Empirical evidence for the formation of the tool	Unclear	Yes	Yes	Yes	No
Sufficient details of the assessment					
The actual in-strument	Yes	Yes	Yes	Yes	Yes
Time of as-sessment	No	Yes	Yes	Partially yes	No
Who are as-sessed	Yes	Yes	Yes	Yes	Yes
The way the assessment is carried out	No	Partially yes	Yes	No	No
Reasonable length	Yes	Yes	Yes	Yes	Yes
Training	Unclear	Unclear	Yes	Yes	Unclear
Cut-off properly decided and clearly specified	Partially yes	Yes	Yes	No	Partially yes
Practical evalua-tion	Yes, but not avail-able	Yes	Yes	No	Yes, but poor quality

None of the assessments described are to be used as screening instruments for child maltreatment. Like many reviews done in the past, no screening instrument has been precise enough. It is unethical and stigmatising to label a family high risk of child abuse. This may antagonise parents and put them off working with health and social service professionals. Therefore, screening is better used to identify children and families in need for more support rather than high risk of child abuse. It is essential that the intervention component is well designed to address problems identified in assessment, such as poor parenting. Classifications of families should be positive (e.g. high priority or low priority for services) rather than negative (e.g. high risk or low risk for maltreatment).

Professionals who are against needs assessment argue that risk factors are no more relevant to the ability to parent than the lack of self-belief. However, in order to target parents who have a low self-esteem related to child care, a needs assessment is required that objectively identifies the needs of the child, the parents' capacity to meet the needs of the child and the social and environmental influences that affect that capacity (e.g. violent partner, social isolation, poverty). However, it is unethical to screen without a

planned intervention strategy to ameliorate undesirable outcomes for child health and development. Therefore, it is important to categorise risk factors into those are static and historical (e.g. history of abuse as a child) and that cannot be changed from those that are dynamic and may be responsive to intervention (e.g. depression, alcohol and drug abuse). Effective intervention can therefore change high need families into low need families by treating those dynamic factors amenable to change.

There is a need for further good quality cohort studies to determine and improve the accuracy of the proposed assessments. Due to the low base rate of child maltreatment in general population, the total sample in a prospective cohort study has to be very large in order to detect enough maltreating families for the analysis to be meaningful. One way to overcome the low base rate is to use case control design where there can be an equal size of cases and controls. However, this does not bypass the need to investigate how an instrument performs on the real population.

Furthermore, the sample selection and attrition should be better recorded and analysed. More details on the sample demographics and characteristics should be provided, as they may explain the findings and give a clearer indication of the applicability of the tool. To ensure objective assessment of the outcomes, the outcome assessors should be blind to the risk/need status of the participants and collect data independently of those who administer the instrument. Finally, the follow up period should be long enough and clearly reported in the publication. It is often overlooked that the predictive accuracy of an assessment only applies to the period of follow up in the study. It is important for agencies to commission studies on assessment to allow the investigation to go on long enough in order to determine long term risk.

References

Agathonos-Georgopoulou, H. & Browne, K. D. (1997). The prediction of child maltreatment in Greek families. *Child Abuse & Neglect, 21,* 721-735.

Ammaniti, M., Speranza, A. M., Tambelli, R., Muscetta, S., Lucarelli, L., Vismara, L. et al. (2006). A prevention and promotion intervention program in the field of mother-infant relationship. *Infant Mental Health Journal, 27,* 70-90.

APHO & Department of Health (2006). *Health Profiles for Surrey 2006* London: Department of Health.

APHO & Department of Health (2007a). *Birmingham Health Profiles 2007* London: Department of Health.

APHO & Department of Health (2007b). *Manchester Health Profiles 2007* London: Department of Health.

APHO & Department of Health (2007c). *Southend on Sea Health Profiles 2007* London: Department of Health.

Appleton, J. (1997). Establishing the validity and reliability of clinical practice guidelines used to identify families requiring increased health visiting support. *Public Health, 111,* 107-113.

Appleton, J. (1995). Health visitor assessment of vulnerability. *Health Visitor., -*31.

Appleton, J. (1994). The role of the health visitor in identifying and working with vulnerable families in relation to child protection: a review of the literature. *Journal of Advanced Nursing, 20,* 167-175.

Appleton, J. V. & Cowley, S. (2004). The guideline contradiction: health visitors' use of formal guidelines for identifying and assessing families in need. *International Journal of Nursing Studies, 41,* 785-797.

Barlow, J., Coren, E., & Stewart-Brown, S. S. B. (2003). Parent-training programmes for improving maternal psychosocial health. *Cochrane Database of Systematic Reviews, Issue 4.*

Barlow, J., Davis, H., McIntosh, E., Jarrett, P., Mockford, C., & Stewart-Brown, S. (2007). Role of home visiting in improving parenting and health in families at risk of abuse and neglect: Results of a multicentre randomised controlled trial and economic evaluation. *Archives of Disease in Childhood, 92,* 229-233.

Barlow, J., Stewart-Brown, S., & Callaghan, H. (2003). Working in partnership: the development of a home visiting service for vulnerable families. *Child Abuse Review, 12,* 172-189.

Brockington, I. F., Oates, J., George, S., Turner, D., Vostanis, P., Sullivan, M. et al. (2001). A screening questionnaire for mother-infant bonding disorders. *Archives of Women's Mental Health, 3,* 133-140.

Browne, K. (1995). Preventing child maltreatment through community nursing. *Journal of Advanced Nursing, 21,* 57-63.

Browne, K. D. (2007). A community health approach to the assessment of infants and their families. In London: CPHVA.

Browne, K. D., Douglas, J., Hamilton-Giachritsis, C., & Hegarty, J. (2006). *A Community Health Approach to the Assessment of Infants and their Parents: The CARE Programme.* Chichester: Wiley.

Browne, K. D., Hamilton, C. E., Hegarty, J., & Blissett, J. (2000). Identifying need and protecting children through community nursing home visits. *Representing Children, 13,* 111-123.

Browne, K. D. & Herbert, M. (1997). *Preventing Family Violence.* Chichester: Wiley.

Browne, K. D. & Saqi, S. (1988). Approaches to screening for child abuse and neglect. In K.Browne, C.Davies, & P.Stratton (Eds.), *Early Prediction and Prevention of Child Abuse* (pp. 57-88). Chichester: Wiley.

Burns, J. (1985). The vulnerability factor. *Nursing Times, 81,* 27-30.

Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T. et al. (2004). Parent-child interaction therapy with physically abusive parents: efficacy for reducing future abuse reports. *Journal of Consulting & Clinical Psychology, 72,* 500-510.

Coren, E. & Barlow, J. (2001). Individual and group-based parenting programmes for improving psychosocial outcomes for teenage parents and their children. *Cochrane Database of Systematic Reviews, Issue 3.*

Cowley, S. & Houston, A. M. (2003). A structured health needs assessment tool: acceptability and effectiveness for health visiting. *Journal of Advanced Nursing, 43,* 82-92.

Cox, A. D. (1997). Preventing child abuse: a review of community-based projects II: issues arising from reviews and future directions. *Child Abuse Review, 7,* 30-43.

Davies, H., Day, C., & Bidmead, C. (2002a). *Working in Partnership with Parents: The Parent Advisor Approach.* London: The Psychological Corporation.

Davies, H., Day, C., & Bidmead, C. (2002b). *The Parent Advisor Training Manual.* London: The Psychological Corporation.

Diareme, S., Tsiantis, J., & Tsitoura, S. (1997). Cross-cultural validation of the Child Abuse Potential Inventory in Greece: a preliminary study. *Child Abuse & Neglect, 21,* 1067-1079.

Dixon, L., Browne, K., & Hamilton-Giachritsis, C. (2005). Risk factors of parents abused as children: a mediational analysis of the intergenerational continuity of child maltreatment (Part I). *Journal of Child Psychology & Psychiatry & Allied Disciplines, 46,* 47-57.

Elkan, R., Kendrick, D., Hewitt, M., Robinson, J. J. A., Tolley, K., Blair, M. et al. (2000). The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature. *Health Technology Assessment, 4,* 1-338.

Elkan, R., Robinson, J., Williams, D., & Blair, M. (2001). Universal vs. selective services: the case of British health visiting. *Journal of Advanced Nursing, 33,* 113-119.

Farrington, D. P. (1995). The 12Th Jack-Tizard-Memorial-Lecture - the Development of Offending and Antisocial-Behavior from Childhood - Key Findings from the Cambridge Study in Delinquent Development. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36,* 929-964.

Gordon, R. R. (1977). Predicting child abuse. *British Medical Journal, 1,* 841.

Gray, J. D., Cutler, C. A., Dean, J. G., & Kempe, C. H. (1977). Prediction and prevention of child abuse and neglect. *Child Abuse & Neglect, 1,* 45-58.

Grietens, H., De Haene, L., & Uyttebroek, K. (2007). Cross-cultural validation of the child abuse potential inventory in Belgium (Flanders): Relations with demographic characteristics and parenting problems. *Journal of Family Violence, 22,* 223-229.

Grietens, H., Geeraert, L., & Hellinckx, W. (2004). A scale for home visiting nurses to identify risks of physical abuse and neglect among mothers with newborn infants. *Child Abuse & Neglect, 28,* 321-337.

Guterman, N. B. (1999). Enrollment strategies in early home visitation to prevent physical child abuse and neglect and the "universal versus targeted" debate: a meta-analysis of population-based and screening-based programs. *Child Abuse & Neglect, 23,* 863-890.

Hamilton, C. E. & Browne, K. D. (2002). Predicting physical maltreatment. In K.D.Browne, H.Hanks, P.Stratton, & C.E.Hamilton (Eds.), *Early Prediction and Prevention of Child Abuse: A Handbook* (pp. 41-56). Chichester: Wiley.

Hamilton, C. E. & Browne, K. D. (1999). Recurrent maltreatment during childhood: a survey of referrals to police child protection units in England. *Child Maltreatment, 4,* 275-286.

Hegarty, J. (2000b). *The CARE Programme (Child Assessment Rating and Evaluation) Training Pack.* Southend on Sea, Essex: Southend Community Care Services (NHS) Trust.

Hegarty, J. (2000a). *The CARE Programme (Child Assessment Rating and Evaluation). Assessment Procedures Manual for Health Visitors.* Southend on Sea, Essex: Southend Community Care Services (NHS) Trust.

Johnson, C. (1985). Identifying children at risk: a system for health visitors. *Health Visitor, 58,* 195-196.

Johnson, Z., Howell, F., & Molloy, B. (1993). Community mothers' programme: randomised controlled trial of non-professional intervention in parenting. *BMJ.306(6890):1449-52*.

Kendrick, D., Barlow, J., Hampshire, A., Polnay, L., & Stewart, B. S. (2007). Parenting interventions for the prevention of unintentional injuries in childhood. *Cochrane Database of Systematic Reviews, Issue 4.*

Lealman, G. T., Haigh, D., Phillips, J. M., Stone, J., & Ord-Smith, C. (1983). Prediction and prevention of child abuse - an empty hope? *The Lancet, 25,* 1423-1424.

Lynch, M. & Roberts, J. (1978). Early alerting signs. In A.W.Franklin (Ed.), *Child Abuse: Prediction, Prevention and Follow-up* (pp. 28-38). NY: Longman.

Milner, J. S. (1989). Applications of the Child Abuse Potential Inventory. *Journal of Clinical Psychology., 45,* 450-454.

Milner, J. S. (1990). *An interpretive manual for the Child Abuse Potential Inventory.* Webster, NC: Psytec.

Milner, J. S. (1994). Assessing physical child abuse risk: The child abuse potential inventory. *Clinical Psychology Review, 14,* 547-583.

Milner, J. S., Charlesworth, J. R., Gold, R. G., Gold, S. R., & Friesen, M. R. (1988). Convergent validity of the Child Abuse Potential Inventory. *Journal of Clinical Psychology., 44,* 281-285.

Milner, J. S., Gold, R. G., Ayoub, C., & Jacewitz, M. M. (1984). Predictive validity of the child abuse potential inventory. *Journal of Consulting & Clinical Psychology, 52,* 879-884.

Milner, J. S., Gold, R. G., & Wimberley, R. C. (1986). Prediction and explanation of child abuse: cross-validation of the child abuse potential inventory. *Journal of Consulting & Clinical Psychology., 54,* 865-866.

Milner, J. S. & Wimberley, R. C. (1979). An inventory for the identification of child abusers. *Journal of Clinical Psychology., 35,* 95-100.

Molloy, B. (2002). *Still Going Strong: A Tracer Study of the Community Mothers Programme, Dublin, Ireland. Early Childhood Development: Practice and Reflections. Following Footsteps.* The Hague: Bernard van Leer Foundation.

Olds, D. L., Eckenrode, J., Henderson, C. R., Jr., Kitzman, H., Powers, J., Cole, R. et al. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect. Fifteen-year follow-up of a randomized trial. *JAMA., 278,* 637-643.

Olds, D. L., Henderson, C. R., Jr., Chamberlin, R., & Tatelbaum, R. (1986). Preventing child abuse and neglect: a randomized trial of nurse home visitation. *Pediatrics, 78,* 65-78.

Olds, D. L. (2007). Preventing crime with prenatal and infancy support of parents: The Nurse-Family Partnership. [References]. Ref Type: Generic

Pecnik, N. & Ajdukovic, M. (1995). The Child Abuse Potential Inventory: cross-validation in Croatia. *Psychological Reports, 76,* 979-985.

Peters, R. & Barlow, J. (2003). Systematic review of instruments designed to predict child maltreatment during the antenatal and postnatal periods. *Child Abuse Review, 12,* 416-439.

Pinheiro, P. S. (2006). *United Nations Secretary-General's Study on Violence against Children* NY: UNICEF.

Puura, K., Davis, H., Papadopoulou, K., Tsiantis, J., Ispanovic-Radjkovic, V., Rudic, N. et al. (2002). The Europea Early Promotion Project. *Infant Mental Health Journal, 23,* 606-624.

Rowe, A. & Carey, L. (2004). The effect of population-based health needs assessment on health visitor practice. *Primary Health Care Research and Development, 5,* 179-186.

Rowe, A., McClelland, A., & Billingham, K. (2001). *Community Health Needs Assessment: An introductory guide for the family health nurse in Europe* Copenhagen: WHO Regional Office for Europe.

Tsiantis, J., Smith, M., Dragona, T., & Cox, A. (2000). Early mental health promotion in children through primary health care services: A multi-centre implementation. *International Journal of Mental Helath Promotion, 2,* 5-17.

World Health Organisation (2002). *World Report on Violence and Health* Geneva: WHO.