Risk assessment in child welfare

An issues paper

DoCS
NSW Department of Community Services

New South Wales Government

An issues paper

Risk assessment in child welfare

Authors

Angela White
Peter Walsh

Produced by

www.community.nsw.gov.au

Contents

1.   Introduction

Child protection practice involves the making of decisions regarding the care, protection and welfare of children. These decisions include assessing the potential level of risk of abuse and/or neglect a child is exposed to in their environment, the services that might reduce this risk and the point at which it becomes necessary for a child to be removed from the care of their family.

A key challenge in child protection services is the identification of effective tools and models that assist caseworkers, managers and organisations to ensure that such decisions are based on the current state of knowledge and research not only surrounding child protection practice, but of human decision-making processes. A particular concern is the identification of common human errors in decision making and ways that these may be avoided, especially as such errors in a child protection context can prove fatal.

Risk assessment is a key component of child protection policy and practice, and informs decisions at all levels of child protection service delivery. As a result, there has been considerable interest in the application of risk assessment technology in child protection practice from the outset. At the same time, there has been intense and at times heated debate regarding the most effective methods for assessing risks and the nature of the risk that needs to be assessed.

More recently, this debate has been escalated into what has been described in the literature as the 'risk assessment wars' (Bauman, Law, Sheets & Graham, 2005, 2006; Johnson 2006a, 2006b). Essentially, this so-called war comes down to a consideration of what are the best methods or approaches to assess and 'predict' abuse and/or the recurrence of abuse of children in child welfare cases. Various iterations of this 'war' have debated the relative merits of clinical judgement, consensus or actuarial models for assessing risk. For child protection services, the debate can appear to be about selecting the 'best' model of assessing risk, implying that the selection of this 'best' model or tool will automatically improve child protection decision-making.

For child protection services, this debate is not an academic one. It is a daily concern for child welfare practitioners who are faced with the constant reality of decision-making over the extent to which they should intervene in the lives of families in order to protect the welfare of children where there may be concerns for their safety.

However, the most recent debate on risk assessment has begun to move away from a search for the ultimate tool to a recognition that whilst some tools more accurately classify risk, this does not rule out the need for other approaches (consensus and clinical judgement) in working out what services will help to ameliorate risk as well as help to engage the family with these services.

This debate accepts that there will always be some inaccuracy associated with risk assessment tools. The literature on human decision making processes in child protection clearly spells out that it is crucial to ensure that information about families needs to be gained from a range of sources by a variety of methods to enhance the accuracy of the decision-making process.

Purpose of this paper

This paper aims to explore the different approaches to the assessment of risk in terms of the potential contributions of each set of tools and/or approaches to decision-making in child welfare.

A number of recent reviews have explored both the issues associated with the implementation of risk assessment systems (Rycus & Hughes, 2003) and have compared a range of risk assessment instruments used in child protection services (D'Andrade, Benton & Austin, 2005; Stewart & Thompson, 2004). This paper does not aim to replicate this work.

The paper starts with looking at the rationale for the use of risk assessment instruments in terms of their purpose and as a way of dealing with problems in decision-making in a child protection context. Following this, an overview of risk assessment approaches is provided in terms of two models of risk assessment – consensus-based instruments and actuarial-based instruments. An important aspect in this discussion is the role of clinical judgement and this is also discussed. Finally, the paper outlines a number of outstanding problems and issues associated with risk assessment with a particular focus on issues of accuracy of instruments and implementation problems.

2. Rationale for the use of risk assessment instruments

Assessment in child protection generally involves two distinct processes: assessing the level of risk of future harm for particular children and a contextual assessment of child and family functioning to inform casework decisions and service planning (Shlonksy & Wagner, 2005). However, these processes are also related. Effective risk assessment is dependent on workers gaining a holistic, ecological and empirically based assessment of the family. This includes an understanding of the conditions that brought the family into the child protection system, each individual family member's history, the systems of which the family is a part, and the strengths and resources the family already possesses (Cash, 2001).

The difficulties involved in making clinical decisions about the risk of future harm in specific cases has led to the development of formal risk assessment measures or instruments to overcome the shortfalls of unassisted clinical judgement. Cicchinelli (1995:3) notes that among the most salient deficiencies that encouraged the development of risk assessment models in child welfare were:

- the lack of a rational basis for making decisions regarding the future of abused and neglected children and their families

- subjective and inconsistent decision-making across cases

- ineffective intervention options and inefficient resource allocation

- a lack of agency accountability

- insufficient worker training and on-the-job support.

Whereas clinical judgement draws on factors such as intuitive reasoning, previous experience or 'practice wisdom' and case study, formal risk assessment models identify specific characteristics for examination during the process of risk assessment. These models are categorised as being either consensus-based or actuarial and they are discussed in more detail in Section 3 of this paper.

2.1 Purpose of risk assessment instruments in child welfare

The primary purpose of using risk assessment instruments in child welfare is to determine the severity of abuse or neglect in order to intervene and protect children from future harm (Cash, 2001). Other reasons for their use include: operationalising the concept of risk, thereby providing a systematic and reliable approach to the assessment of risk (Doueck et al, 1993); substantiating a case (Wald & Woolverton, 1990); determining the severity of the abuse or neglect and prioritising cases (English & Pecora, 1994); providing a way to structure documentation throughout the history of a case (Pecora, 1991; Wald & Woolverton 1990); determining the type of services needed; determining critical points in the case such as child removal or case closure (Wald & Woolverton, 1990); and as an organisational risk management strategy (Gambrill & Shlonsky, 2001; Cicchinelli 1995, Tomison, 1999, Wells, 1994).

Despite the many benefits of risk assessment instruments, a number of problems have been identified. For example, Pecora (1991) argues that risk assessment instruments should not be used to substantiate a case because they are generally developed to predict future behaviours, not past. Others argue that risk assessment instruments do not intend to predict future child maltreatment, rather they aim to classify families according to their level of risk with the aim of informing agencies as to the level of services that family is likely to require (Shlonsky & Wagner, 2005). Cicchinelli (1995) notes that it is just such discrepancies of purpose that continue to fuel the controversy over the adequacy and best use of the currently available models.

Although the literature at times wistfully hopes for a tool that can serve all these purposes, there is consensus that no one tool will be capable of doing so (Schene, 1996; English, 1996; Shlonsky & Wagner, 2005) and that achieving a sufficient degree of consensus in the field on factors for inclusion in such an instrument is, if possible, far into the future (Wald & Wolverton, 1990; McDonald & Marks, 1991; Arad, 2001). A key issue is that it will not be possible to eliminate error rates (Daniel, 1978; Caldwell et. al., 1988; Anglin, 2002). As Marzouki (1997 as cited in Goddard, Saunders, Stanley & Tucci, 1999, p. 260) states, "no model of cause can embody the complexity of the reality", a consideration that has led to a concern that reliance on a single instrument could result in legal liability for agencies and/or practitioners (Curran, 1995; Goddard, et al., 1999).

2.2   Problems with decision-making in child protection assessment

A significant issue in child protection risk assessment is the inconsistencies in decision-making processes of child protection service workers. In a study on errors of reasoning in child protection work in the UK, Munro (1999) points out that complete accuracy in child protection assessment is an unrealistic expectation. However, the results from this study examining 45 child abuse inquiry reports pointed out that 'in some of these cases, it appeared that on the evidence available, professionals drew the wrong conclusions and that the subsequent tragedies and distress could have been averted' (Munro, 1999: 745-746).

An underlying problem is two different approaches to human reasoning: analytical and intuitive. Analytical reasoning is described as 'a step-by-step, conscious, logically defensible process' as opposed to intuitive reasoning which is 'a cognitive process that somehow produces an answer, solution or idea without the use of a conscious, logically defensible, step-by-step process' (Hammond, 1996: 60). In child protection practice, many professionals rely heavily on intuitive skills (Munro, 1999) despite the evidence that "intuition is a hazard, a process not to be trusted, not only because it is inherently flawed by 'biases' but because the person who resorts to it is innocently and sometimes arrogantly overconfident when employing it" (Hammond, 1996: 88).

Stewart & Thompson (2004) summarise some of the literature on human decision-making relating to practitioners' prediction of risk. Four biases have been identified:

- practitioner's tendency to under use base rates when predicting events that are uncommon (which leads to a tendency to overestimate the occurrence of an event)

- confirmatory biases often prevent practitioners from considering evidence impartially (which leads to a tendency to search for evidence consistent with the conclusion they believe to be correct)

- illusory correlations have been found to influence clinical predictions (which leads to a tendency to see two events as being related when they are not, or are related to a lesser extent)

- practitioners tend to place too much importance on the unique characteristics of a case (which leads to a tendency to believe that similar cases are quite different and that unique characteristics are better predictors than those that are more common).

The literature on human reasoning and decision-making indicates that personal judgement is often influenced by contextual factors such as the representativeness of the case, the availability or vividness of information and the presumed relevance of the available information to the decision being made (Cicchinelli 1995). Munro (1999: 754) found that most determinations of risk were based on a limited range of data, often with the most memorable cases (those that aroused emotion or were most recent) factoring into the assessment of risk more than the 'dull, abstract material in research studies, case records, letters and reports'. Subsequently, even with evidence contrary to the workers initial case disposition, revision of judgement about cases was slow or non-existent.

Decisions were also often faulty due to biased reporting or errors in communication. A critical attitude to evidence was found to correlate with whether or not the new information supported the existing view of the family. A major problem was that professionals were slow to revise their judgements despite a mounting body of evidence against them (Munro 1999).

Munro (1999) concluded that errors in professional reasoning in child protection work are not random but predictable on the basis of research on how people intuitively simplify reasoning processes in making complex judgements. As a result, "analytical tools are needed to supplement intuitive skills and shift practice reasoning along the continuum towards the analytical end" (Munro, 1999: 754). Aids to reasoning need to be developed that recognise the central role of intuitive reasoning but offer methods for checking intuitive judgements more rigorously and systematically. Hence, risk assessment instruments have the potential to improve practitioner reasoning and decision-making.

3. Overview of risk assessment approaches

A widely used approach in the literature to categorise risk assessment approaches is to distinguish between models on the basis of how the specific instrument was developed – whether derived through consensus of experts or empirically developed (actuarial).

3.1 Consensus-based models

Consensus based instruments are those with items that are included based on the clinical judgement of experts (Baird & Wagner, 2000; Cash, 2001; Gambrill & Shlonsky, 2000). Consensus risk assessment models are 'compiled by "experts" who may draw upon previous research findings, clinical experience, or a combination of both, but an empirical study is usually not conducted to validate the assessment' (Shlonsky & Wagner, 2005: 410).

A well known example of a consensus-based instrument is the Washington Risk Assessment Matrix (WRAM). The WRAM tool has 37 items covering seven domains: child characteristics, severity of abuse/neglect, chronicity of abuse/neglect, caretaker characteristics, caretaker/child relationship, socio economic factors and perpetrator access. Practitioners assess and rate the level of risk on a five point scale and as a result families are categorised into risk levels. In a recent review of this tool (D'Andrade et. al., 2005), both predictive validity and convergent validity showed mixed results and inter-rater reliability was poor.

Poor reliability and validity are not uncommon for these type of tools. As a result, Cash (2001: 818) excluded consensus based instruments from his discussion on the science of risk assessment on the basis that 'a risk assessment instrument not based on empirical findings is nothing more than practice wisdom arranged neatly on a form'.

3.2 Actuarial-based models

Actuarial instruments, on the other hand, are developed by employing empirical research procedures to identify a set of risk factors with a strong statistical relationship to the behavioural outcome of interest to the predictive enterprise (Shlonsky & Wagner, 2005). These are then weighted and combined to form an assessment tool that optimally classifies families or individuals according to the 'risk' that they will exhibit the behaviour. Actuarial instrument items included are empirically identified by studying child protection cases and abuse/neglect outcomes in a specified geographical area (Baird & Wagner, 2000; Cash, 2001; Gambrill & Shlonsky, 2000; Schwalbe, 2004). As a result, caution needs to be employed in utilising these approaches in different jurisdictions without appropriate validation studies.

A well known example of an actuarial instrument is the Structured Decision Making (SDM) risk assessment tool developed by the Children's Research Centre. The instrument includes two subscales of ten items each – one for risk of neglect and one for risk of physical or sexual abuse. Based on the subscale scores, families are classified into a low, moderate, high or very high category. The tool allows for policy or discretionary over-ride to increase the risk rating. Recent reviews of this tool (D'Andrade et. al., 2005; Stewart & Thompson, 2004) indicate high predictive validity and inter-rater reliability for this instrument and it is currently being used by child protection authorities in South Australia and Queensland, as well as a number of states in the US[1].

---

[1] The approach used in New South Wales is a guided practice risk assessment model that includes the collection and analysis of information and the exercise of professional judgement. Apart from Queensland and South Australia, all other Australian jurisdictions use a professional judgement model for risk assessment. For a national comparison of child protection systems in Australia see Bromfield and Higgins (2005) available at http://www.aifs.gov.au/nch/issues/issues22.html

3.3   Comparison of approaches

There is a strong body of research indicating that actuarial approaches are superior to either clinical judgement or consensus-based approaches in assessment of risk (Ruscio, 1998; Dawes, Faust & Meehl, 1989; Leschied et. al., 2003), particularly in relation to the classification of families at risk for child maltreatment (DePanfilis & Zuravin, 2001; Baird et al., 2000).

For example, Grove and Meehl (1996) report on a meta-analysis of 136 studies comparing clinical with statistical predication across a range of health and behavioural settings. Approximately 95% of the studies found actuarial approaches to be equal to or superior to clinical judgement. Even a small set of actuarial factors generally provides a better prediction of outcomes than does the use of clinical judgement.

Consensus-based approaches have been criticised for several reasons (Rycus & Hughes, 2003; Wald & Woolverton, 1990; McDonald & Marks, 1991):

- poor conceptualisation of measures

- inconsistency in type and number of variables used across consensus-based instruments

- use of the same variables to predict all types of maltreatment despite differences in the dynamics between neglect, physical abuse and sexual abuse

- lack of a focus on recurrence of maltreatment.

At the same time, however, actuarial methods are far from infallible (Ruscio 1998; Dawes et. al., 1989). These models have considerably less accuracy in determining which moderate risk families are most likely to become high risk, or which families are at risk for tragic outcomes such as child death (Sanders & Goddard, 1998). There is also recognition that there has been little work on whether the factors which predict abuse are the same as those predicting re-abuse (Lyons et al., 1996; Wald & Woolverton, 1990).

There are a number of limitations to the actuarial approach that are summarised by Hart (1998):

- actuarial approaches tend to focus assessment on a limited number of factors, thus ignoring potentially crucial case-specific, idiosyncratic factors

- there is a tendency to focus on relatively static factors that are immutable, therefore leading to passive predictions

- actuarial approaches may exclude crucial risk factors on the basis they have not been proven empirically, even though they may be logical (for example, homicidal threats)

- actuarial approaches tend to be optimised to predict a specific outcome, over a specific time period in a specific population, leading to non-optimal decisions when applied in different settings (Gottfredsen & Gottfredsen 1996 as cited in Doyle & Dolan, 2002).

3.4   The role of clinical judgement

Most research on risk assessment acknowledges that the use of any kind of risk assessment instrument, actuarial or consensus-based, requires good clinical skills (Doueck, English, DePanfilis & Moote 1993). For example, the SDM™ actuarial risk assessment tool requires clinical judgement to score and allows for a clinical override based on family characteristics or dynamics that are likely to affect risk but are not on the actuarial instrument (D'Andrade et. al., 2005: 29).

As Ereth et. al. (2003: 3 cited in D'Andrade et al 2005) note:

> *a caseworker can sense things that an actuarial instrument would ignore or could not employ… Many characteristics of human subjects simply cannot be quantified empirically and actuarial models cannot easily account for rare event.*

Therefore, clinical judgement can never be eliminated from any risk assessment process. Many researchers in child welfare stress that instruments for risk and safety assessment should be understood as decision aids to enhance or expand upon clinical judgement rather than as a competing approach (Shlonsky & Wagner, 2005; Ereth et. al., 2003; Fuller et. al., 2001; Munro, 1999).

Although actuarial methods are more likely to provide accurate information, there are still a number of difficulties associated with their use as indicated above. It is important to emphasise that risk assessment instruments are in essence risk classification tools rather than abuse *prediction* tools (Shlonsky & Wagner, 2005; DePanfilis & Zuravin, 2001; Baird & Wagner, 2000). Silver and Miller (2002: 147) note that 'actuarial risk assessment is more useful in labelling individuals than it is in understanding and resolving their problems'. As Wilczynski (1997) points out, they are better at distinguishing between groups than individuals. Thus, instead of predicting what will occur, classification of greater or lesser degree of risk simply informs practitioners and agencies about which cases are more likely than others to be high-risk. As a result, the professional judgement of workers is still crucial.

3.5   Combination of approaches

Doyle and Dolan (2002) note that practitioners are concerned with the clinical reality of assessing and managing risk rather than the research task of prediction. Consequently, they propose a 'third generation' approach, referred to as 'empirically validated, structured decision-making' (Douglas et. al., 1999) or 'structured clinical judgement' (Hart, 1998). This approach attempts to bridge the gap between the scientific, actuarial approach and the clinical practice of risk assessment. The emphasis is on developing evidence-based guidelines or frameworks that promote systemisation and consistency yet are flexible enough to account for case specific influences and the contexts in which assessments are conducted. Such instruments can promote transparency and accountability yet encourage use of professional discretion and are based on sound scientific knowledge but are practically relevant (Hart, 1998).

Saunders and Goddard (1998) stress that no risk assessment checklist will include every possible risk factor or address the issues of 'volatile combinations' of circumstances that can precipitate child maltreatment. Therefore, findings from risk assessment instruments should be used with caution and other methods of assessment (for example, observation) should be employed (Shlonsky & Wagner, 2005; Baird & Wagner, 2000).

In practice, assessment of families is generally dependent upon a variety of formal and informal activities and observations, including record reviews, personal contacts with the client, collateral contacts with law enforcement, school, medical and social service personnel and, in many instances, consultation with colleagues and supervisors (Baird et. al., 1999). The application of clinical judgement is seen as assisting in the detection of those factors which are of such low frequency that actuarial methods are unlikely to assist with their detection.

In the late 1990s in Australia, Tomison (1999) and Saunders and Goddard (1998) argued that the best approach to risk assessment is a combination of clinical judgement and actuarial methods. Although there has been concern that clinical judgements alone can produce 'self-fulfilling' prophecies biasing the outcome (Munro 2004a; Dawes et. al., 1989), the most recent literature on risk assessment instruments highlights the need to refocus on the contribution of clinical judgement to child welfare decision-making (Shlonsky & Wagner 2005).

Shlonsky and Wagner (2005) argue that risk assessments should be combined with an objective assessment of family strengths and needs by the clinician and the information from both assessments translated into the choice of a set of effective service interventions. Consequently, the use of risk assessment instruments is not seen as replacing the need for professional and well-trained staff. The importance of adequate and appropriate worker training is stressed. For example, Doueck and colleagues (1993) point out that without good quality control and worker supervision the risk assessment system can be used to support potentially poor decisions. They stress that:

> these models cannot replace a competently trained staff – in fact, all … models require a staff that is trained and knowledgeable in human growth and development, parenting practices, the causes and effects of mistreatment and family dynamics.

This issue is explored further in the next section.

4.   Problems and issues in risk assessment

This section explores a number of problems and issues associated with the use of risk assessment tools and approaches in the child welfare context. Four broad inter-related issues are identified. First, there are significant definitional issues involved in risk assessment. Finally, there are a number of implementation issues that are crucial to risk assessment approaches.

4.1   Definitional issues

Numerous analyses of risk assessment instruments have identified the lack of agreed definitions of risk as a fundamental problem, affecting both the empirical validation of these instruments and their implementation in the field (Wald & Woolverton, 1990; McDonald & Marks, 1991; Goddard et al, 1999; Rycus & Hughes, 2003).

There is no present risk assessment system that defines, in quantitative terms, 'high', 'medium' or 'low risk' (Wald & Woolverton, 1990). For example, we do not know if classifying a family as 'high risk' means there is a 10%, 30% or an 80% probability that a family will, in fact, re-abuse children. The same is true of the term 'low risk'. The best that can be said for existing instruments is that they are able to rank cases, more or less accurately, along a risk continuum, without specifying how close the case is to either end of the continuum, or how much difference there is between cases with different rankings.

4.2   Accuracy of risk assessment instruments and the problem of false negatives and false positives

The risk assessment process is critical to determining how child protection agencies respond to cases. There is evidence that the introduction of formal risk-assessment instruments into child welfare practice, both consensus and actuarial, lead to a more consistent response (Baird et. al., 1999; Doueck et. al., 1993). As noted earlier, actuarial models are more accurate than clinical judgement or consensus models (Leschied et. al., 2003).

However, no method of assessment will have 100% reliability and the rarer the event (such as fatal child assault) the more difficult it will be to predict without unacceptably high levels of 'false positives'[2] (Ruscio 1998). Current estimates of the accuracy of actuarial instruments in predicting child maltreatment range from around 70% to 80% (Leschied et al 2003; Doueck et al 1993) compared with 64% for clinical decision making. Anglin (2002: 243) indicates that accuracy of such tools is, 'at best, not likely to exceed 80%'. English and Graham (2000) also observe that the ability to predict new referrals, though easier both in terms of obtaining a large sample of cases to study and for making accurate predictions, does not necessarily correspond to an ability to predict *recurrences*.

The accuracy of any risk assessment instrument is determined by three variables (Munro, 2004):

1. the sensitivity of the instrument (how many high risk families are correctly identified – true positives)

2. the specificity of the instrument (how many low risk families are correctly identified – true negatives)

3. the base rate or prevalence of the problem being measured (child maltreatment).

---

[2]   False positives are those instances when a child is unnecessarily identified as being at risk of maltreatment. False negatives are those instances when a child who is at risk of maltreatment is not identified.

It is the last of these variables that is particularly important. The rarer the problem being assessed, the harder it is to develop an instrument with a high level of accuracy. The small numbers of child abuse cases that occur within the population (less than 1%) (AIHW 2003) and the even smaller number of fatal child abuse cases (around 0.001%) (CDRT 2004) make it almost impossible to generate accurate risk assessment tools.

As a result, risk assessment approaches can result in levels of false negatives and false positives that may be unacceptable to the community (Ansay & Perkins 2001). These approaches are on the one hand over-inclusive and generate a high number of false positives and on the other are insufficiently sensitive and generate a high number of false negatives, especially since the problem is rare (Wilczynski 1997). Parton, Thorpe and Wattam (1997: 59-60) illustrate this in the following way:

> *If the criterion to predict abuse is 80 per cent accurate (and no techniques have developed such high levels of accuracy) and the rate of abuse in the population is 10 in 100, then if 1,000 people/families are screened, 80 out of the hundred will be correctly identified while 20 will be missed (false negative). However, we will also incorrectly identify 20 per cent, or 200 of the original 1,000 screened. Thus for every 80 correctly predicted we will have 20 false negatives and 200 false positives.*

There are considerable costs associated with the problem of false negatives and false positives, including the fiscal costs of assessing and supporting those families who were not at risk for maltreatment as well as responding to those families who abused their children but were not identified as being at risk. There are also costs for the families who are labelled as being at risk but will not abuse their children, such as the potential for such labelling to create legal, civil and psychological problems (Caldwell, Bogat & Davidson 1988). The experience of investigation can be traumatic for families and may not result in service provision (Cleaver & Freeman 1995).

4.3   The move from prediction to classification in risk assessment

Much of the discussion on risk assessment both in the literature and in practice discussions tends to lack clarity around the distinction between 'predictive' and 'actuarial' models of risk assessment (Schene, 1996). Early risk assessment instruments focussed on attempting to *predict* the cases that would result in subsequent maltreatment and so led to a search for the appropriate factors, appropriate weighting of factors and increasing knowledge about what constitutes the 'riskier' combinations or interactions of factors. Later developments in risk assessment included a distinction between these predictive models and actuarial models which aimed to help staff identify and *classify* cases by risk of future maltreatment. In other words, the focus is not on prediction per se, but on grouping certain types of cases as more likely than other groups of cases to result in subsequent maltreatment and the risk assessment protocol is judged on its ability to sort out the high-risk cases (Schene, 1996).

Shlonsky and Wagner (2005) highlight the inherent difficulties in prediction measures which use a traditional binary (yes/no) decision schema. As noted above, this leads to high levels of false positives and false negatives. Given that human behaviour is extremely difficult to predict, a more useful schema is to *classify* cases into varying levels of risk (for example, low, moderate, high and very high). This does not provide certainty that a particular child will be maltreated. Rather, the risk classification indicates there is a greater likelihood of maltreatment for one case than for others. Once a classification has been made, providing an estimate of the risk of future harm, clinical judgement is then required to determine the nature of further intervention:

> *Once the risk assessment is scored, clinical judgement must still be employed to make the final decision whether to offer services or close the case. At this pivotal transition, actuarial estimation is integrated with, and supplemented by, clinical judgement* (Shlonsky & Wagner, 2005: 417-418).

The same holds true for consensus based risk assessments. In this way, a risk classification is not a substitute for the use of sound professional judgement by practitioners but rather an aid for improving clinical decisions. Grove and Meehl (1996) highlight this in terms of the 'broken leg' problem. An actuarial assessment could provide an estimate of the probability of an individual undertaking a particular behaviour (such as going to the movies on a specific evening), based on an analysis of a range of relevant variables. However, if the individual has a broken leg and is immobilised in a hip cast, the actuarial assessment will be wrong. This kind of information is only accessible if the individual is interviewed or observed by the clinician.

Clinical judgement is needed to identify and clarify specific problem behaviours and the context in which they occur.

4.4 Implementation issues in risk assessment

One of the key issues associated with risk assessment approaches of whatever kind is the need to address implementation issues. This section explores a number of these issues identified in the literature including the need for training and supervision, the impact of organisational factors, the emergence of risk management by child protection authorities and the costs of risk management.

Training and supervision

The literature is clear in identifying that training and supervision are crucial in addressing implementation issues. For example, implementation studies on risk assessment systems indicates that risk assessment demands quality in education, training and supervision, as well as vigilance and persistence on the part of administrators hoping to use it (Doueck, English, DePanfilis & Moore, 1993; English & Pecora, 1994). Scott (2003) discusses findings from a study which found that the diversity of child abuse work creates anxiety amongst practitioners and that poor case supervision can confound long term plans, lead to poor support for parents and carers and result in anger and hostility from family members. Limited training, inadequate supervision and the lack of a supportive peer group were found to reduce awareness of important issues with a bearing on child abuse such as domestic abuse or depressive illness. Surbeck (2003), following a study into racial partiality of child welfare assessments of attachment, concludes that there is a need for training around attachment, racial differences and self-awareness for child protection service staff.

Organisational context and culture

Assessments made by child protection service workers do not rely solely on factors within the worker, but also on organisational factors. Khoo, Hyvönen and Nygren (2003) point out that how workers respond to families at the point of intake depends on whether an institution adopts a child protection or child welfare stance. A child protection stance involves work being coordinated by a standard set of processes child protection workers follow to define the child protection features of a case. The child welfare stance ensures that professionals in child welfare have a set of standardised skills, decision-making is decentralised and social workers rely on an investigatory process to produce cumulative information with which to assess a child's needs. The authors suggest that the ideal model would combine the best features of the child protection and child welfare orientations, responding to need and ensuring safety.

Schwalbe (2004) notes that risk assessment may be used by frontline workers to circumvent agency policy or ignored due to doubts about its usefulness (also see Doyle & Dolan, 2002). Results of risk assessments may be skewed to ensure families receive services (English & Graham, 2000).

Lyons and colleagues (1996) note that the available evidence suggests that risk assessment models are being imperfectly implemented, that measures are being completed after the assessment decisions have been made, merely as a means of documenting decisions rather than as a guide to the decision making process itself and finally that they are perceived by some workers as irrelevant to their work with both 'at risk' and maltreating families.

Gambrill and Shlonsky (2001) identify the systemic deficits in child welfare that relate to risk assessment and impact on individual worker decisions. These are:

- poor organisational culture within the agency

- ineffective tracking of mistakes/errors and the corrective action taken to solve them

- lack of available services

- available services that lack effectiveness

- vague contracts between agencies and service providers

- lack of recurring program evaluations.

As Ryan and colleagues (2005) point out, these environmental factors create a propensity for perpetuating risk to clients and reinforce the need for standardised procedures and processes that assist workers in making the best, most reliable decisions about the case.

Risk assessment as a risk management strategy

Risk assessment systems themselves are seen as organisational risk management strategies (Gambrill & Shlonsky, 2001). Cicchinelli (1995) notes that the speed with which the concept of risk assessment and the accompanying models, checklists and continuums have been embraced is not as much an indicator of the quality of the assessment approaches available as it is an indication that child protection service agencies are under ever increasing scrutiny. Tomison (1999) points out that media reports have resulted in a "moral panic" about child protection and the channelling of resources into investigation and detection at the expense of supportive and preventative services. Child welfare agencies must have established criteria against which the appropriateness of decisions made can be assessed (Cicchinelli, 1995).

Munro (2004b) identifies that one way in which agencies protect themselves is by 'blame-prevention engineering' and one strategy for doing this is 'protocolisation': the organisation introduces more and more formal procedures to guide practice so that they create a 'correct' way to deal with a case. Then, if a tragedy occurs, they can claim the defence of 'due diligence' and show that their employees followed these correct procedures in working on the case. Munro notes that this defence seems to be operating in the UK where recent inquiries into the deaths of children known to child protection agencies appear to focus more on whether procedures were followed than whether competent professional judgements and decisions were made.

Costs of risk assessment

The costs associated with risk assessment, particularly the impact on families, is one of the areas that featured strongly in early discussions but less so more recently. Given that the impact of the low base rate phenomena on the capacity to accurately predict risk means that current risk assessment procedures provide a relatively small increase in program efficiency, the costs of using them must be evaluated carefully. Two kinds of costs are involved: the cost to the agency conducting or underwriting the assessment and the cost to the person forced to undergo the assessment procedure (Caldwell et al, 1988).

5. Conclusion

In the period 1990-2000, the literature surrounding risk assessment tended to be polarised into a debate on the relative merits of actuarial, consensus and clinical judgement approaches to risk assessment. Whatever the approach adopted, it was noted (and still remains the case) that risk assessment is hampered by:

- problems with definitions of harm and risk (Wald & Woolverton,1990)

- problems with implementation (Munro, 1999)

- level of error (Anglin, 2002)

- lack of empirical studies into the actuarial tools available. Most studies were implementation studies (Camasso & Jagannathan, 2000) and some of the key studies on SDM, for example, were conducted by researchers now employed by CRC (Baird, Wagner, Healy & Johnson, 1999; Baird & Wagner, 2000).

At the same time, the literature drew on actuarial studies (such as Dawe et al, 1989) to demonstrate the capacity of actuarial approaches to improve consistency and accountability in decision making across child protective services (Ruscio, 1998). Concern over clinical judgement approaches focussed on human information processing errors, particularly the idea that personal judgement is often influenced by contextual factors such as the representativeness of the case, the availability or vividness of the information and the presumed relevance of the available information to the decision being made (Cicchinelli, 1995; Munro 1999).

However, actuarial risk assessment tools will not necessarily assist in predicting which children within the population known to child protective services are at risk of severe injury or death. The rarity of such events poses challenges for actuarial instruments and it has been argued that, rather than to try to identify these cases, effort is better directed towards prevention initiatives for all families as such strategies are more likely to reach these families than more targeted initiatives (Wilczinski, 1997; Reder & Duncan, 1999).

Tomison (1999) and Sanders and Goddard (1998) critiqued the 'either/or' debate between actuarial and consensus/clinical judgement approaches, suggesting that the different approaches to risk assessment were complementary rather than mutually exclusive.

There is consensus in the literature that, whatever the approach adopted, its efficacy is dependent on skilled, well-trained and supported staff (Doueck et al, 1993).

References

AIHW, (2004). *Child Protection Australia 2002-3*. Australian Institute for Health and Welfare: Canberra.

Anglin, J.P., (2002). Risk, Well-being and paramountcy in child protection: the need for transformation. *Child and Youth Care Forum.* 31:233-255.

Ansay, S.J. & Perkins, D.F., (2001). Integrating family visitation and risk evaluation: a practical bonding model for decision makers. *Family Relations. 50*(3): 220-229.

Arad, B.D., (2001). Parental features and quality of life in the decision to remove children at risk from the home. *Child Abuse and Neglect. 25*(1): 47-64.

Baird, C. & Wagner, D., (2000). The relative validity of actuarial and consensus based risk assessment systems. *Children and Youth Services Review. 22*(11/12): 839-871.

Baird, C., Wagner, D., Healy, T. & Johnson, K., (1999). Risk assessment in child protective services: consensus and actuarial model reliability. *Child Welfare. 78*(6): 723-748.

Baumann, D.J., Law, J.R., Sheets, J., Reid, G. & Graham, J.C., (2005). Evaluating the effectiveness of actuarial risk assessment models. *Children and Youth Services Review.* 27: 465-490.

Baumann, D.J., Law, J.R., Sheets, J., Reid, G. & Graham, J.C., (2006). Remarks concerning the importance of evaluating actuarial risk assessment models: A rejoinder to Will Johnson. *Children and Youth Services Review.* 28: 715-725.

Bromfield, L. & Higgins, D., (2005). National comparison of child protection systems. *Child Abuse Prevention Issues.* No. 22.

Caldwell, R.A., Bogat, G.A. & Davidson, W.S., (1988). The assessment of child abuse potential and the prevention of child abuse and neglect: a policy analysis. *American Journal of Community Psychology.* 16: 609-624.

Camasso, M.J. & Jagannathan, R., (2000). Modeling the reliability and predictive validity of risk assessment in child protective services. *Children and Youth Services Review. 22*(11&12): 873-896.

Cash, S.J., (2001). Risk assessment in child welfare: the art and science. *Children and Youth Services Review. 23*(11): 811-830.

Child Death Review Team (CDRT) (2004) *Fatal Assault and Neglect of Children and Young People 2003.* NSW Commission for Children and Young People, Sydney.

Cicchinelli, L., (1995). Risk assessment: expectations and realities. *The Apsac Advisor. 8*(4): 3-8.

Cleaver, H. & Freeman, P., (1995). *Parental perspectives in cases of suspected child abuse.* London: HSMO.

Curran, T. F., (1995). Legal issues in the use of CPS risk assessment instruments. *The APSAC Advisor, 8*(4), 15-21.

D'Andrade, A., Benton, A. & Austin, M., (2005). *Risk and Safety Assessment in Child Welfare: Instrument Comparison.* Berkeley: Bay Area Social Services Consortium.

Daniel, J.H., Newberger, E.H., Reed, R.B. & Kotelchuk, M., (1978). Child abuse screening: implications of the limited predictive power of abuse discriminants from a controlled family study of pediatric social illness. *Child Abuse & Neglect.* 2: 247-259.

Dawes, R.M., Faust, D. & Meehl, P.E., (1989). Clinical vs Actuarial Judgement. *Science. 243*(4899): 1668-1674.

DePanfilis, D. & Zuravin, S.J., (2001). Assessing risk to determine the need for services. *Children and Youth Services review. 23*(1): 3-20.

Doueck, H.J., English, D.J., DePanfilis, D. & Moore, G.T., (1993). Decisionmaking in child protective services: a comparison of selected risk-assessment systems. *Child Welfare. 72*(5): 441-453.

Douglas, K., Cox, D. & Webster, C., (1999). Violence risk assessment: science and practice. *Legal and Criminological Psychology.* 4: 149-184.

Doyle, M. & Dolan, M., (2002). Violence risk assessment: combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing*. 9: 649-657.

English, D., (1996). Promise and Reality of Risk Assessment. *Protecting Children, 12*(2), 9-13.

English, D.J. & Graham, J.C., (2000). An examination of relationships between Children's Protective Services social worker assessment of risk and independent LONGSCAN measures of risk constructs. *Children and Youth Services Review. 22*(11/12): 897-933.

English, D.J. & Pecora, P.J., (1994). Risk assessment as a practice method in child protective services. *Child Welfare. 73*(5): 451-473.

Gambrill, E. & Shlonsky, A., (2001). The need for comprehensive risk management systems in child welfare. *Children and Youth Services Review. 23*(1): 79-107.

Gambrill, E. & Shlonsky, A., (2000). Risk assessment in context. *Children and Youth Services Review, 22*(11/12), 813-837.

Goddard, C. R., Saunders, B. J., Stanley, J. R., & Tucci, J., (1999). Structured Risk Assessment Procedures: Instruments of Abuse? *Child Abuse Review, 8*(4), 251-263.

Grove, W. & Meehl, P., (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychology, Public Policy and Law.* 2: 293-323.

Hammond, K. R., (1996). *Human Judgement and Social Policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press.

Hart, S.D., (1998). The role of psychopathy in assessing risk for violence: conceptual and methodological issues. *Legal and Criminological Psychology.* 3: 121-137.

Johnson, W., (2006a). The risk assessment wars: A commentary response to Evaluating the effectiveness of actuarial risk assessment models by Donald Baumann, J. Randolph Law, James Sheets, Grant Reid and J. Christopher Graham, Children and Youth Services Review, 27, pp. 465-490. *Children and Youth Services Review, 28*, 704-714.

Johnson, W., (2006b). Post-battle skirmish in the risk assessment wars: Rebuttal to the response of Baumann and colleagues to criticism of their paper, "Evaluating the effectiveness of actuarial risk assessment models". *Children and Youth Services Review*, in press.

Khoo, E.G., Hyvönen, U. & Nygren, L., (2003). Gatekeeping in child welfare: A comparative study of intake decision making by social workers in Canada and Sweden. *Child Welfare. 82*(5): 507-525.

Leschied, A.W., Chiodo, D., Whitehead, P.C., Hurley, D. & Marshall, L., (2003). The empirical basis of risk assessment in child welfare: the accuracy of risk assessment and clinical judgement. *Child Welfare. 82*(5): 527-540.

Lyons, P., Doueck, H.J. & Wodarski, J.S., (1996). Risk assessment for child protective services: a review of the empirical literature on instrument performance. *Social Work Research. 20*(3): 143-156.

McDonald, T., & Marks, J., (1991). A review of risk factors assessed in child protective services. *Social Service Review, 65*(March 1991), 113-132.

Munro, E., (1999). Common errors of reasoning in child protection work. *Child Abuse & Neglect. 23*(8): 745-758.

Munro, E., (2004a). A simpler way to understand the results of risk assessment instruments. *Children and Youth Services Review.* 26: 873-883.

Munro, E., (2004b). The impact of child abuse inquiries since 1990. In Stanley J & Manthorpe J (Eds.) *The Age of the Inquiry: Learning and blaming in health and social care*. Ch. 4 pp. 75-91. London: Routledge.

Munro, E., (2005). Improving practice: child welfare as a systems problem. *Children and Youth Services Review.* 27:375-391.

Palmiter, D.J., (2004). A survey of the assessment practices of child and adolescent clinicians. *American Journal of Orthopsychiatry. 74*(2):122-128.

Parton, N., Thorpe, D. & Wattam, C., (1997). *Child Protection: Risk and the moral order.* London: Macmillan.

Pecora, P.J., (1991). Investigating allegations of child maltreatment: the strengths and limitations of current risk assessment systems. In M. Robin (Ed.), *Assessing reports of child maltreatment: The problem of false allegations* (pp. 73-92). Binghamton NY: Haworth Press.

Reder, P., & Duncan, S., (1999). Auditing Mental Health Aspects of Child Protection. *Child Abuse Review, 8*(3), 147-151.

Ruscio, J., (1998). Information integration in child welfare cases: an introduction to statistical decision making. *Child Welfare. 3*(2): 143-156.

Ryan, S., Wiles, D., Cash, S. & Siebert, C., (2005). Risk assessments: empirically supported or values driven? *Children and Youth Services Review.* 27: 213-225.

Rycus, J.S. & Hughes, R.C., (2003). *Issues in risk assessment in child protective services: policy white paper.* Ohio: Center for Child Welfare Policy.

Schene, P., (1996). Risk Assessment Roundtables: A Ten-Year Perspective. *Protecting Children, 12*(2), 4-8.

Schwalbe, C., (2004). Re-visioning risk assessment for human service decision making. *Children & Youth Services Review, 26*(6), 561-576.

Scott, L., (2003). Vulnerability, need and significant harm: an analysis tool. *Community Practitioner. 76*(12): 468-473.

Shlonsky, A., & Wagner, D., (2005). The next step: Integrating actuarial risk assessment and clinical judgment into an evidence-based practice framework in CPS case management. *Children & Youth Services Review, 27*(4), 409-427.

Silver, E. & Miller, L.L., (2002). A cautionary note on the use of actuarial risk assessment tools for social control. *Crime & Delinquency. 48*(1): 138-161.

Stewart, A. & Thompson, C., (2004). *Comparative Evaluation of Child Protection Assessment Tools.* Queensland: Griffith University.

Surbeck, B.C., (2003). An investigation of racial partiality in child welfare assessments of attachment. *American Journal of Orthopsychiatry. 73*(1): 13-23.

Tomison, A.M., (1999). *Ensuring the protection of children: the role of child protection services in the identification, assessment and treatment of maltreated children.* Keynote address to the NSW Department of Community Services, Entry Parameters for Child Protection Conference, March 20-21, 1999, Sydney.

Wald, M.S. & Woolverton, M., (1990). Risk assessment: the emperor's new clothes? *Child Welfare. 64*(6): 483-511.

Wilczynski, A., (1997). *Child Homicide.* London: Greenwich Medical Media.

Wells, S.J., (1994). Child protective services: research for the future. *Child Welfare.* 73: 431-450.