Roger Almvik

Assessing the Risk of Violence:

Development and Validation of the Brøset Violence Checklist

Thesis for the Degree of Doctor Philosophiae

Trondheim, March 2008

Norwegian University of
Science and Technology
Faculty of Medicine
Department of Neuroscience
# Table of contents

ACKNOWLEDGEMENTS ........................................................................................................ 4

LIST OF PAPERS ............................................................................................................... 5

ABBREVIATIONS .............................................................................................................. 6

1.0 INTRODUCTION ........................................................................................................ 7

2.0 AIMS ......................................................................................................................... 12

3.0 RISK OF VIOLENCE ............................................................................................... 13

3.1 PREDICTING VIOLENT BEHAVIOUR .................................................................... 16

4.0 INPATIENT VIOLENCE IN PSYCHIATRY ............................................................... 19

4.1 FACTORS RELATED TO VIOLENCE IN PSYCHIATRY ....................................... 19

4.1.1 Patient-related factors ....................................................................................... 19

4.1.2 Unit-related factors ........................................................................................... 22

4.1.3 Staff-related factors ........................................................................................... 23

4.1.4 Interactional variables ...................................................................................... 25

4.2 THE PREVALENCE OF INPATIENT AGGRESSION ............................................... 26

4.3 MEASURING AGGRESSION OF PSYCHIATRIC PATIENTS ................................ 30

4.4 RESEARCH ON INPATIENT VIOLENCE IN NORWAY ........................................ 41

5.0 DEVELOPING A TOOL FOR VIOLENCE RISK ASSESSMENT .................. 47

5.1 DEVELOPING THE BRØSET VIOLENCE CHECKLIST ....................................... 47

5.2 INTERNATIONAL VALIDATION OF THE BVC .................................................. 49
# METHODS

## 6.1 Methodological Considerations

## 6.2 Statistics

### 6.2.1 Summary

# SUMMARY OF PAPERS

## 7.1 Paper I

## 7.2 Paper II

## 7.3 Paper III

## 7.4 Paper IV

## 7.5 Paper V

# DISCUSSIONS

## 8.1 Clinical Implications and Future Research

# REFERENCES

# ERRATA

# APPENDIX
Acknowledgements

This dissertation is the result of many efforts and contributions alongside my own. My friends and collaborators throughout many years at the forensic department and research centre at Brøset, the University Hospital of Northern Norway and the European Violence in Psychiatry Research Group: huge thanks to you all. Thanks also to the nursing staff at all participating study centres in Trondheim, Tromsø and Oslo for their efforts and skills in gathering data. I could not have done this without you! I am especially indebted to my supervisors, Professor Kirsten Rasmussen and Dr Phil Woods. They have both been very supportive and inspiring as well as very patient with a slow learner……..

I also want to thank my dear colleague and friend, Trond Hatling at SINTEF Helse. His ability to support and encourage is unsurpassed. His positive attitude and enthusiasm have been an invaluable help.

A number of other persons have also contributed to the present study in different ways. I wish to express my deep appreciation to Professor Olav Linaker for initiating this work, to Dr Egil Bjarnar, Dr Jim Aage Nøttestad and Dr Tom Palmstierna for continuous support and encouragement, and to Professor Henk Nijman for inviting me in to his work and world of knowledge.

Finally I want to thank my partner Astrid for her support and caring during these years. The atmosphere she and our two children Martin and Kristin have created at home has been of major importance for this project.
List of papers

The Brøset Violence Checklist (BVC) and the prediction of inpatient violence: some preliminary results. Psychiatric Care, 5 (6), 208-211


Fifteen years of research with the Staff Observation Aggression Scale; A review. Acta Psychiatrica Scandinavica, 111, 12–21


Assessing risk for imminent violence in the elderly; the Brøset Violence Checklist. International Journal of Geriatric Psychiatry, 22, 862-867
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>Area Under Curve.</td>
</tr>
<tr>
<td>BVC</td>
<td>Brøset Violence Checklist.</td>
</tr>
<tr>
<td>HCR-20</td>
<td>Historical, Clinical and Risk management, 20 items.</td>
</tr>
<tr>
<td>PCL</td>
<td>Psychopathy Checklist.</td>
</tr>
<tr>
<td>PICU</td>
<td>Psychiatric Intensive Care Unit.</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver Operating Characteristics.</td>
</tr>
<tr>
<td>SCU</td>
<td>Special Care Unit (for patients with dementia).</td>
</tr>
<tr>
<td>TBS</td>
<td>Terbeschikkingstelling; Dutch judicial measure.</td>
</tr>
<tr>
<td>SOAS-R</td>
<td>Staff Observation Aggression Scale – Revised.</td>
</tr>
<tr>
<td>WAS</td>
<td>Ward Atmosphere Scale.</td>
</tr>
</tbody>
</table>
1.0 Introduction

Risk assessment in general and the risk of violence in particular have received more and more attention in mental health care during the last couple of decades (Holdsworth, Collis, & Allott, 1999). This is clearly illustrated by the media attention every time a mentally ill person is involved in serious crime. Despite both awareness of and competence on the topic, it remains a complex issue, influenced to some degree by who defines the risk (McClelland, 1995).

The Oxford Dictionary defines the term “risk” as “the probability of a bad consequence” (The Oxford Dictionary, 2005). Prins (1996) describes it as the likelihood that a particular adverse event will recur. Overall, most definitions of risk focus on outcome and probability.

In the research literature, the term “violence” has often been used interchangeably with assault and aggression, and with a wide range of definitions. However, in most cases one or more of the following components are included: behaviour damaging to individuals or property; attitude, moods or gestures experienced as threatening or intimidating; sexual harassment and stalking (Wood, 1987 p. 17, Sandberg, 1998). A common definition of violence often used in mental health settings in Europe is “any verbal, non-verbal, or physical behaviour that was threatening (to self, others or property), or physical behaviour that actually did harm to self, others, or property” (Morrison, 1990). Consequently, this definition was used in the studies reported here in this thesis.
Studies of psychiatric patients involve a wide range of indicators of aggression, e.g. criminal records, self-report of violent acts, physical attacks on staff, verbal abuse and threatening behaviour, self-harm, and damage to property.

Assessment of the risk of violence has become more and more important in mental health care, and it is widely accepted that mental health professionals should have knowledge and experience of risk assessment. Anderson (2004) indicates two reasons for this. First, failing to acknowledge or ignoring the risk to those working with mentally ill patients leaves staff unprepared for exposure to violence. This can lead to situations in which staff are less willing to work with violent patients, which may in turn result in a lack of resources for this group of patients. Second, Anderson points out that society has become more aware of the possible link between mental illness and violence.

In forensic psychiatry, the focus has commonly been on long-term prediction, while in clinical psychiatry the focus is on short- and medium-term prediction (Lidz, 1993, Johnson, 2000). Short-term prediction is seen to be the more problematic, due partly to the absence of relevant measurement tools (Soliman & Reza, 2001).

Despite substantial progress in the treatment of mental health problems (as a result of improved neuroleptics and more effective psychosocial interventions), the frequency of violent incidents (violence towards self or others) among psychiatric inpatients is on the increase (Nijman, Muris, Merckelbach, Palmstierna, Wistedt, Vos, van Rixtel, & Allertz, 1999). Benjaminsen and Kjærbo (1997) reported that more than 90% of doctors and nurses have been exposed to violence during their career in mental health care in Denmark, with nurses being more frequent victims of patient-to-staff violence.
Several studies of working conditions confirm the high prevalence and the variation between professions. These studies indicate that exposure to inpatient violence has a marked impact on the well-being of staff and on professional stability in these groups; it threatens the safety and well-being of staff members and fellow patients (Benjaminsen & Kjærbo 1997, Arnetz, Arnetz, & Soderman, 1998, Bowers, Jarrett, Clark, Kiyimba, & McFarlane, 1999).

Nursing staff in particular appear to be at risk of being assaulted by their patients (Carmel & Hunter, 1989; Gerberich et al., 2004; Nijman, Merckelbach, Allertz, & a Campo, 1997).

Violent episodes harm both staff and patients (Hyde & Harrower-Wilson, 1995). They strain the economy of the health care system, for instance through increased levels of sickness absence (Hunter & Carmel, 1992; Viitasara & Menckel, 2002), they also harm the relationship between staff and the patient as well as the therapeutic atmosphere on the wards (Cutting, 1997).

Nursing staff are essential in the delivery of high quality care, but when exposed to violence they often experience a persistent reduction in quality of life (Arnetz et al., 1998; Whittington & Wykes, 1989) and they may become suspicious and hostile in their working situation (Bowers et al., 1999).

In a recent survey among 148 psychiatric nurses in East London (Nijman, Bowers, Oud, & Jansen, 2005), almost one in six nurses (16%) claimed to have experienced severe physical violence in the workplace during the past year. Just over one in five nurses (22%) confided that they had called in sick due to workplace violence.
It is well documented that staff involved in violent incidents may develop post-traumatic stress disorder (Richter & Berger, 2006; Whittington & Wykes, 1992). In other words, violence from psychiatric patients not only has considerable physical and psychological consequences for the staff, but is also very likely to have substantial financial implications for the health service. This is endorsed by a much cited study by Hunter and Carmel (1992) in which a total of 134 serious injuries were reported in a single year at a 973-bed forensic psychiatric hospital. The average cost per injury was conservatively estimated to be $5,719, making the total annual loss $766,290. Furthermore, data from the EU indicates a significant correlation between health-related workplace absence and exposure to violence at work (European Foundation for the Improvement of Living and Working, 2000, p.5).

A number of studies indicate that the prevalence of workplace violence has increased over the last few years and is still increasing (Gerberich et al., 2004). The National Health Services Executive in the United Kingdom (2002) found that 65,000 violent incidents occurred against staff within that year. The average number of incidents within mental health and learning disability services in NHS trusts was more than three times the average for the total sample, giving a rate of 24 violent incidents a month per 1000 staff within mental health/learning disability services. In a survey of incidents involving violence towards staff caring for elderly people, Åstrøm and colleagues (Astrom, Bucht, Eisemann, Norberg, & Saveman, 2002) found that 40% of the respondents had been exposed to violence during the preceding year. Nursing-home staff reported the highest incidence, with 55% affected.
A study on 10 acute wards in Norway found that 65% of nursing staff had been involved in a violent episode during the past 12 months, and more than 80% regarded violence as a "daily risk" in their work (Berg, Sveipe & Hoy 1994).

Figures such as these clearly illustrate the magnitude of the problem, but there appears to be a lack of well-designed studies, using standardised instruments, of violence in psychiatric institutions in Norway. Hence we know little about this phenomenon in terms of the frequency, causes or consequences of violence. The obvious limitations associated with the use of survey instruments, such as recall and selection bias, indicate the necessity for continuous monitoring of aggressive episodes on psychiatric wards as they happen. Reliable and time-efficient methods of recording aggression should be in place at psychiatric wards to enable the collection of complete and accurate information on the incidence of aggression in the wards, and the magnitude of the problem. Continuous monitoring of incidents on the ward may also be helpful in detecting typical precursors and triggers of violent behaviour.

Presumably, the range of options for intervention will increase with more knowledge about when and why patients are most likely to engage in aggressive behaviour. The high prevalence of aggressive incidents and the financial as well as the human costs associated with inpatient violence clearly indicate the need for studies and methods that can predict and hence prevent workplace violence. To reduce the problem of inpatient violence in psychiatric settings, we need good assessment methods using both clinical and systematic information in order to guide interventions and prevent recurrence of the violence.
2.0 Aims

The overall aim of this thesis was to validate the Brøset Violence Checklist (BVC) in various mental health settings. This was achieved through testing the psychometric properties of the BVC (i.e inter-rater reliability and predictive validity) alongside a well validated dependent measure; the Staff Observation Aggression Scale, revised version (SOAS-R). Studies allowed some description of the prevalence and nature of violent incidents in acute, psychogeriatric and special care for dementia settings.

Furthermore, this thesis will describe and discuss the problem of inpatient violence in mental health care as well as violence risk assessment issues, and will provide a brief introduction to vital issues in these problem areas. Different methods of measuring violence will be addressed and the process of developing the BVC will be described.
3.0 Risk of violence

“A structured risk assessment acts as an aide-mémoire, making sure that we collect all the relevant information… Best practice, therefore, is to use the data as the basis for a clinical team meeting. The teams do the work of evaluation and planning, once the structured instrument has ensured that the necessary information is to hand. The outcome is, of course, only as good as the clinical team. If staff lack training or experience, they end up by being overwhelmed by information they are unable to use”

(Maden, 2003, p.201)

Among researchers and clinicians with an interest in various types of violence, there have traditionally been two approaches to risk assessment. Most common has been the unstructured clinical risk assessment, also referred to as professional assessment or first-generation risk assessment. This method was predominantly in use from the 1960s to the early 1980s, and was characterised by a lack of rules or exact methods; it depended solely on the professional experience of the assessor. This method has been extensively criticised, primarily because it was rarely, if ever, possible to validate the assessments in retrospect and because the processes leading to the conclusions were obscure. In his review of the “first generation” of violence prediction research, Monahan (1981) found that psychologists and psychiatrists were accurate in no more than one out of three predictions of violent behaviour among mentally ill inpatients with a history of violence. Rice, Harris and Quinsey (2002) did
not, however, interpret this to mean that clinicians’ predictions were worse than chance, but rather that they reflected over-cautiousness.

The other approach, *actuarial risk assessment* – also referred to as the second generation of risk assessment – came into use during the late 1980s. Unlike clinical assessment, the actuarial approach is strictly based on defined rules and on data that previous research has shown to be highly correlated with violent behaviour. The critics of this method have focused on the exclusion of clinical experience and judgement, and regard it as too “mechanical”, relying heavily on measures that are essentially static in nature. Levander (2000) pointed out that this method did not identify many of the most severe crime and violence recidivists and that risk assessment should include observations on paranoid behaviour, extreme lack of empathy, perversions, dissociations and psychopathy.

The more recent approach, *structured professional risk assessment*, was introduced in the mid-1990s and labelled third-generation risk assessment (Belfrage & Fransson, 2000). This approach is characterised by the use of checklists with items based mostly on empirical data. Clinicians use the information from the checklists to underpin their judgement, then “filter” the information on the basis of their experience and personal knowledge about the actual case. This method encourages the assessor to take into account historical individual factors as well as specific clinical and situational factors, combined with future risk factors, treatment plans and security measures, often referred to as a contextual, dynamic continuous model (Borum, Fein, Vossekuil & Berglund, 1999).
The Historical-Clinical-Risk Management Scheme, 20 items; HCR-20 (Webster, 1997) is a typical and very successful representative of the latter approach. This instrument has been widely applied in the assessment of future violence risk among people with mental disorders in forensic settings. The instrument is in use worldwide; numerous studies support both its accuracy and its inter-rater reliability (Belfrage, Fransson, & Strand, 2000). A major issue among the developers and users of this instrument has also been to divert the focus away from the purely predictive aspect towards risk management and prevention.

Recently, Monahan et al. (2005) have developed an actuarial model of risk assessment for people with mental disorders, the multiple Iterative Classification Tree (ICT). Showing promising results in assisting discharge planning for acutely hospitalised patients, the ICT is mainly concerned with community violence and will not be discussed further here.

In spite of general agreement among clinicians and researchers on the superiority of the structured clinical approach to risk assessment over the previous methods, there is equally general agreement that the results are far from perfect. In addition, the focus has been on long-term prediction. Other characteristics of these instruments are that they require extensive information about the patient and are time-consuming to use, rendering them less useful to nursing staff in need of a quick and easy-to-use instrument. There seems to be a clear need for a short and accurate measure of the prevalence and nature of violent incidents in mental health care, and several attempts have been made to meet this need. Risk assessment and measurement based on the two most commonly used methods – self-report questionnaires and observation scales – are discussed in a later chapter.
3.1 Predicting violent behaviour

The most frequently used methods for assessing aggression of psychiatric patients can be roughly divided into self-rating and observer aggression scales (Bech, 1994). A well-known self-rating questionnaire for measuring hostility and anger is the Buss-Durkee Hostility Inventory (BDHI; Buss-Durkee, 1957), which has a long history. Research over the years on the psychometric qualities of BDHI items has led to adapted versions of the BDHI, such as the Aggression Questionnaire (AQ; Buss and Perry, 1992).

On the basis of the literature, Bjørkly (1995) noted about the BDHI that “(…) its predictive value for adult psychiatric patients has not been convincing so far” (p. 49). Indeed, much of the research on the psychometric properties of aggression self-report measures has traditionally been done in normal subjects, such as psychology students (Yudofsky, Silver, Jackson, Endicott, & Williams, 1986), but this seems to be changing rapidly (Novaco & Taylor, 2004). One of the problems with aggression self-reports may be that patients with severe psychiatric disorders, such as antisocial personality disorder or schizophrenia, lack insight into their own role in initiating conflicts. Furthermore, self-reports of aggressive behaviour rely heavily on the honesty of respondents about their tendency to become angry and behave aggressively. With regard to such self-reporting in general, Yudofsky and colleagues (1986) asserted that “many patients are not angry between aggressive episodes, and do not reliably recall or admit to past violent events” (p. 35).

Especially in forensic psychiatric samples, in which release from custody may be linked to the psychiatric condition, the inclination to provide socially desirable
answers may pose great problems for the validity of self-reported symptoms, including aggressiveness. This seems to be illustrated by a Dutch study by Hornsveld and colleagues (2004). These researchers found that forensic patients sentenced to one of the Dutch “TBS” hospitals reported lower aggression and hostility scores than forensic outpatients outside the TBS system. On the basis of their violent and criminal histories, the opposite would have been expected. Possibly, outpatients’ answers were more honest because high hostility scores in this group could not lead to longer admission periods, whereas for the sample of incarcerated TBS patients a judge has to decide every two years on continuation or termination of the TBS sentence. In other words, apart from a lack of insight into one’s own behaviour, a problem with aggression self-reports may be that certain subgroups have a strong interest in playing down their impulse control problems.

On the other hand, several recent studies found evidence for the predictive validity of self-reports on feelings of anger and aggressiveness, even in forensic psychiatric patient samples. To give an example: Novaco and Taylor (2004) measured aggressiveness and anger among 129 male forensic patients with intellectual disabilities using modified versions of the Novaco Anger Scale (NAS; Novaco, 1994) and the Spielberger State-Trait Anger Expression Inventory (STAXI), (Spielberger, 1996). In this study, significant correlations were found between the STAXI “Trait Anger” and “Anger Expression” subscales on the one hand, and hospital physical

---

1 Dutch criminal law states that a defendant who, at the time of the alleged crime, was affected by a mental defect or disorder may receive what is called a “disposal to be involuntarily admitted to a forensic psychiatric hospital on behalf of the state”, in Dutch “maatregel van terbeschikkingstelling” or TBS.
assaults on the other (correlations of 0.34 and 0.37, respectively). The correlation between hospital assaults and the NAS total score was 0.43 (Novaco & Taylor, 2004).

In addition, aggression self-report questionnaires are generally designed to assess tendencies of individuals to react in a hostile or an angry way; they rarely include questions about discrete incidents. They are thus not designed to provide information on the prevalence of aggressive incidents in specific wards or institutions. For this reason, much of the scientific literature on the prevalence as well as on the prevention of aggressive behaviour on psychiatric wards has relied on the use of aggression observation scales. Ward staff generally use observation scales to record the aggressive behaviour of psychiatric patients after aggressive incidents have occurred. In contrast to aggression self-report questionnaires, most observer aggression instruments are intended for measuring discrete incidents.
4.0 Inpatient violence in psychiatry

4.1 Factors related to violence in psychiatry

There is a wealth of literature focusing on factors related to violence among psychiatric inpatients and a large body of research has identified a number of variables which seem to be associated with violent behaviour. In this chapter, however, only a small proportion of this literature will be presented and discussed to illustrate the diversity and magnitude of this research area.

Early research on violence in psychiatry focused to a large degree on patients and patient-related factors. However, Steinert (in Richter & Whittington, 2006) claims that “there is no evidence that environmental and interactional influences are less important in their contribution to the origin of violence” (p.113). In the author’s view, recent research has taken a much broader perspective, indicating that intrainstitutional violence is related to a combination of patient-, staff- and environment-related factors. In a review paper, Johnson (2004) claimed that violence in inpatient psychiatric units could be grouped into four categories: patient-related, staff-related, unit-related, and interactional variables.

4.1.1 Patient-related factors

Two of the patient-related factors involved in violent episodes among inpatients in psychiatry are diagnosis and age. Age seems to be one of the most common demographic variables studied. Several studies conclude that younger
patients are more likely to become violent (Chou, Lu, & Mao, 2002; Omerov, Edman, & Wistedt, 2002), while other studies reveal that elderly patients are more likely to be aggressive (Owen, Tarantello, Jones, & Tennant, 1998a); still others found no significant differences (Apter, Plutchik, & van Praag, 1993), so findings are somewhat inconsistent.

A variety of psychiatric diagnoses have been reported to be associated with violence. The most consistent finding seems to be the correlation between schizophrenia and aggression (Ehmann et al., 2001; Flannery & Walker, 2001; McNiel & Binder, 1994) while others indicate a relationship between mania and subsequent aggression (McNiel & Binder, 1994, 1995). Several studies have concluded that mild mental retardation (Powell, Caan, & Crowe, 1994), substance abuse (Flannery & Walker 2001) and personality disorder (Raja, Azzoni, & Lubich, 1997) are significantly related to violent behaviour, but there is no consistent pattern in research results linking a particular diagnosis to violent episodes.

The gender issue has also been studied in the context of inpatient violence. Many studies did not find any gender differences in patients involved in violent behaviour (e.g. Chou et al., 2002), others found higher prevalence in men (e.g. Steinert, Hermer, & Faust, 1996), while still others found female inpatients to be more violent (e.g. Flannery & Walker 2001). Krakowski & Czobor (2004) found that women with positive psychotic symptoms were more likely to be violent and violence in men was more frequent when related to substance abuse. Again, results seem to be contradictory and a clear conclusion is not possible.

A previous history of violence is a strong and robust predictor of violent inpatient behaviour. Many studies have confirmed this, and according to Steinert (Steinert,
2002) no published studies have actually questioned this. Other historical variables that are related to violence are previous hospitalisation and the total length of hospitalisation (Chang & Lee, 2004).
4.1.2 Unit-related factors

In her review, Johnson (2004) also found several studies on unit-related variables associated with violence. These included staffing ratios, ward rules/house rules, census, patient mix and location of incidents. As mentioned in the introduction to this chapter, she too noticed that studies of unit-related factors seem to introduce the notion that the problem of violence in inpatient units is not necessarily related to the staff or patients, but is rather an interaction among staff, patients and the environment.

A number of studies show that the frequency of violence during the day varies. Some studies show that most severe incidents occur in the morning (e.g. Chang & Lee, 2004), while others find that most incidents take place at mealtimes (Lanza, Kayne, Hicks, & Milner, 1994). Chou et al. (2002) also found mealtimes critical, specifically between 12:00 and 14:00 and similarly between 17:00 to 19:00. Others found no pattern related to violence and time of the day (Omerov et al. 2002; Owen, Tarantello, Jones, & Tennant, 1998b). These varied findings could easily lead to a discussion focused on ward-specific cultures/factors and structures of the day rather than on global theories on the relation between time of day and violence in institutional mental health care.

The location of the incidents has also been studied. Nijman et al. (1997) found that living rooms and dayrooms were the places within the ward that were associated with the highest risk of violent episodes. Chou et al. (2002) found that most incidents take place in patients’ bedrooms, in front of the nursing station and in the hallways. Again, the variation in findings could reflect local cultures and differences.
In developing the Ward Atmosphere Scale (WAS), Moos (1974) focused on conceptualisation and measurement of the psychosocial milieus in the psychiatric wards. The WAS emphasizes the interactions in treatment milieus as perceived by patients and staff. In Norway, Friis and colleagues have conducted a number of important studies on environmental factors, and Friis (1991) concludes that: “There is a lack of empirical studies investigating the interaction between environmental factors and violence”. A further illustration of this is the book by Krøvel, Rund and Rør (eds.; 1997) on violence and psychiatry: several chapters are based on research on patient factors, while the only chapter focusing on environmental factors is based solely on clinical experience. Both clearly point out the need for more research into environmental and/or unit-related factors for inpatient violence.

4.1.3 Staff-related factors

The effect of staffing levels on patient-to-staff violence is still unclear and little systematic research has been done. Both Chou et al. (2002) and Lanza et al. (1994) revealed that higher patient-to-staff ratios were associated with more violence in the unit. Owen et al. (1998a) found that increasing the number of staff on a ward was correlated with a higher prevalence of violence. The authors put forward several reasons for their findings, which again highlights the complexity of this problem; the effect of staffing levels related to violence remains unclear. According to Larkin, Murtagh & Jones (1988), acute wards have a higher level of incidents than other types of hospital wards. Whittington & Wykes (1994), however, claims that there is no clear connection between ward type and violence.
Katz & Kirkland (1990), report a number of distinct differences between what they label as “peaceful” and “violence disposed” units regarding social organisation: routines and ward rules, patient-staff relations, the use of therapeutic procedures and behaviour of the psychotherapist. The “peaceful” units had a more ritualised use of routines and rules and more trusting, calm and predictable patient-to-staff relations, extensive use of therapeutic procedures, well-defined organisation of the staff and a dedicated psychiatrist. The “violence-disposed” units had directly opposite characteristics. Black et al. (1994) point out that an important factor in explaining violence in hospitals is the differences in how incidents are managed. Surprisingly, they found that doctors and psychiatrists felt safest in hospitals with a high frequency of incidents; staff in hospitals with a lower rate of incidents were described as inattentive and poorly trained in how to handle incidents.

Staff-related factors have been studied by a range of researchers, and include staff age and gender, experience and education level.

Benjaminsen and Kjærbo (1997) found no significant relationship to gender or to age in staff; however, they found some evidence that the lower the level of education, the greater the risk of being attacked. On the other hand, Flannery and Walker (2001) found that male staff were more likely to be victims of violence, while Owen et al. (1998b) found that a higher ratio of female to male staff was correlated with increased prevalence of violence in a hospital unit. The same study indicated that the younger (age < 30 years) the staff, the lower the risk of violence, but also concluded that staff without training in either in psychiatry or management of aggression were more likely to be involved in violent incidents.
4.1.4 Interactional variables

Johnson (2004) describes interactional variables as variables related to the relationships among patients and staff on the unit. These relations include interactions, interaction style and staff-patient rapport.

As stated and documented elsewhere in this thesis, nurses and other staff in continuous contact with patients are more likely to be assaulted than, for example, doctors and psychologists who usually spend less time with patients; the more time you spend with the patients, the higher the risk of being assaulted (Flannery & Walker 2001).

Omerov et al. (2002) found particular kinds of staff-patient interactions to be linked with violent behaviour: requesting patients to take medication, refusing to allow the patient to leave the ward or unit and denying some kind of reward. Limit setting is also reported as a typical interaction style associated with violence (e.g. Chou et al. 2002). Units with a majority of nurses who perform and interact in a more restrictive and controlling manner have increased frequency of aggression (Duxbury, 2002).

Although knowledge of factors such as those described above will certainly increase the competence of mental health workers in risk assessment and risk management, many questions are left unanswered. Despite the importance of individual factors, experienced nurses will know that violent episodes do not occur in a context-free situation, and further research is still needed.
4.2 The prevalence of inpatient aggression

There is little published evidence indicating the number of violent incidents that occur in psychiatric wards across Europe. However, evidence from a review of studies in selected EU countries indicates the range of incidents in psychiatric wards to be between 0.4 to 33.2 per patient per year (Nijman, Palmstierna, Almvik, Stolker, 2005). The same study shows that 10-20% of these incidents have physical consequences and 1-5% require somatic treatment. In previous descriptive studies, the prevalence of aggressive behaviour ranges considerably: from as low as 0.15 assaults per bed per year (Fottrell, 1980) to as high as 88.8 incidents per bed per year on a specialized high-security ward (Brizer, Convit, Krakowski, & Volavka, 1987). In a review Bjørkly (1995) estimated that 15 to 30% of hospitalised psychiatric patients are involved in physical assaults.

In individuals with dementia, aggressive behaviours occur in as many as 65%, increasing to 95% if behaviours such as agitation are included (Sourial, McCusker, Cole, & Abrahamowicz, 2001). Furthermore, Ballard et al. (2001) reported a prevalence of 86% in a UK residential care/nursing home sample. Buckley et al. (1990) studied 698 people diagnosed with schizophrenia who had been admitted to a psychiatric hospital in Ireland; 16% of these patients had engaged in violent behaviour since the onset of their illness. In Denmark, Benjaminsen (1991) studied a hospital sample in acute care, finding that the percentage of inpatients who engaged in violent behaviour was 6.2. Like many other studies, Benjaminsen concluded that a small group of patients was responsible for a large percentage of the total number of violent acts in a psychiatric hospital. In the Minnesota Nurses’ study
Gerberich et al. (2004) found an increased rate for both physical and non-physical violence for nursing staff working in a nursing home or a psychiatric/behavioural department.

The above-mentioned differences in findings could be due to a number of reasons: the selection of patients studied or the criteria used to define and measure aggressive behaviour differ considerably across the studies, adding to the inconsistency of the results. Bowers (2000) noted that comparison of aggression frequencies between different wards and hospitals has been severely handicapped by a failure to uniformly express incident rates in the past. One possible way to reduce these differences would be if all researchers used the same instrument to report on violent incidents within their studies. Currently, many researchers use the Staff Observation Scale (SOAS or SOAS-R), thus allowing more meaningful comparison of aggression frequencies across countries and types of wards.

A review of the reported SOAS-R frequencies (Nijman et al., 2005) yielded a median value for all reports of slightly fewer than eight incidents per psychiatric patient per year. This figure of about eight SOAS-R incidents per psychiatric patient per year would mean that on a 15-bed psychiatric ward a SOAS-R form would be completed once every three days. However, the reported annual number of SOAS-R incidents per psychiatric patient still varied considerably across studies, from as low as 0.4 to as high as 59.9 incidents per year, depending on the type of ward and country involved. When the annual number of SOAS-R incidents per patient was studied separately for acute admissions wards taking care of adult patients (n = 38 separate observations) a mean of 9.3 incidents per bed per year was found. Remarkably, several studies on aggression in psychogeriatric samples (conducted in Sweden and the UK) reported a
high prevalence of aggression (more than 15 incidents per patient per year with a maximum of 59.9), but there is some indication that the average severity of the incidents, in terms of physical consequences, is low on such wards. In general, the proportion of incidents leading to physical consequences (e.g., pain, bruises, or welts) ranges from 10 to 20% of the total number of SOAS-R assessments (Nijman, et al., 2005). Even more severe assaults for which victims require somatic treatment constitute about 1 to 5% of all SOAS-R reports.

As could be expected, high SOAS-R-frequencies of incidents were also found on wards providing care to selected groups of high-risk patients (e.g., 40.2 and 29.2 in selections of violent schizophrenic patients in Finland, and 31.2 in a selection of young Dutch high-risk patients who required involuntary admission at a young age). With the Report Form for Aggressive Episodes (REFA), Bjørkly (1996) also found high aggression frequencies on specialized wards, such as a 19-bed Norwegian special secure unit for dangerous psychotic patients. In a 10-year prospective study, 2021 incidents of aggressive behaviour were recorded in total. Equivalent values of rates of aggression per patient per year were: 25.9 (total of aggressive episodes), 13.5 (verbal threats), 6.5 (physical threats) and 5.9 (physical assaults). Four patients accounted for 1558 (77%) of the aggressive episodes that occurred during the study period.

Several SOAS-R studies also reported reductions in aggression reports over time, but these studies covered shorter periods. Among others, Nilsson, Palmstierna & Wistedt (1988) have speculated that this “reduction phenomenon” could be “caused by a learning process from the ordinary nursing staff, who during a study of this kind are forced to systematize their observation of their patients” (p.174). Alternatively, the ‘spontaneous’ decrease in registration of incidents over time could be related to
changes in the way that aggression observation scales are completed as time progresses (Sival, Albronda, Haffmans, Saltet, & Schellekens, 2000, De Niet, Hutschemaekers, & Lendemeijer, 2005).

The preliminary evidence from the SOAS-R review suggested further that there may be differences in prevalence rates in the various European countries. The mean number of incidents per patient per year from the Dutch acute admissions wards, for instance, was high when compared with the mean number of incidents from the other countries (e.g., the UK, Germany, Norway, and Denmark), but the relatively low number of studies per country does not allow for very firm conclusions on this finding.

There is still little unambiguous documentation on reasons for and prevalence of violent behaviour and this can be traced back to the use of non-standardised methods in the research field. However, the growing use of standardised aggression report forms from the scientific literature in psychiatric institutions across Europe is likely to improve the comparability of aggression frequencies between wards, hospitals and countries in the future. Possibly, interesting cross-national differences will emerge when aggression frequencies can be compared on a larger scale.
4.3 Measuring the aggression of psychiatric patients

Despite many years of mental health reforms, including the implementation of a deinstitutionalisation programme and increased focus on outpatient treatment, the problem of violence is still a key issue in mental health care. Reliable and clinically relevant recording of the frequency and manifestations of assaultive behaviour is needed to provide service managers and clinicians with an understanding of the extent and the patterns of aggression among psychiatric inpatients. Factors such as the under-reporting of assaults, incomplete and inconsistent operational definitions, the lack of distinction between major and minor assaults as well as between verbal and physical assaults, and the particular victims of violent behaviour all make it difficult to compile a clear picture of a potentially assaultive psychiatric inpatient. These are important issues to bear in mind when evaluating previous research and when planning future studies. There appears to be a lack of well-designed and uniform studies using standardised instruments on violence in psychiatric institutions, at least in Norway. In order to develop a prevention policy, it is essential to have a reliable picture of the number, nature and severity of aggressive incidents within a care institution or on a ward.

This chapter deals with issues associated with measuring violence in mental health care. A range of different methods for recording incidents as well as different instruments will be described and discussed. Finally, a summary of the various methods and instruments is presented, together with the rationale for using one particular instrument to validate the BVC.
The problem of the under-reporting of incidents of aggression is very important, but it will not be discussed in this chapter, as the issue is dealt with in more detail in paper 3 and paper 4 in this thesis.

The literature indicates that there are substantial differences in the prevalence of aggression in different psychiatric hospitals and wards. Some reviews have explained these differences in terms of the varieties of aggression registration procedures used. It is also suggested that ward staff themselves are part of the interactions that lead to aggressive incidents on psychiatric wards (Whittington & Wykes, 1996), which may affect the objectivity of the ratings. One way to avoid this might be to videotape ward activity (Crowner, Stepcic, Peric, & Czobor, 1994; Nolan & Volavka, 2006), so that conflicts on the ward could be rated by independent observers later on. Due to practical and ethical problems, this approach, however, is not in use in clinical practice.

For the prevention of aggressive behaviour in inpatient facilities, the recording of incidents immediately after they occur may have advantages. For instance, this system may make it possible to investigate the specific circumstances and times that are associated with violent outbursts (Nijman et al., 1997). The recording of aggressive behaviour directly after each incident may also be useful in helping to provide more insight about factors which elicit aggression (Nilsson, Palmstierna, & Wistedt, 1988).
There are many different instruments for recording observations of aggression, which require completion after an incident has occurred. They include:

- Overt Aggression Scale, OAS; (Silver & Yudofsky, 1991; Yudofsky, Silver, Jackson, Endicott, & Williams, 1986);
- The modified version of this instrument, MOAS; (Kay, Wolkenfeld, & Murrill, 1988);
- Staff Observation Aggression Scale, SOAS; (Palmstierna & Wistedt, 1987);
- The revised version of SOAS: SOAS-R (Nijman, Muris, Merckelbach, Palmstierna, Wistedt, Vos, van Rixtel & Allertz, 1999);
- Report Form for Aggressive Episodes; REFA (Bjørkly, 1996)
- Rating Scale for Aggression in the Elderly, the RAGE; (Patel & Hope, 1992)
- Brief Agitation Rating Scale: BARS (Finkel, Lyons, & Anderson, 1993).
- Cohen-Mansfield Agitation Inventory, CMAI (Cohen-Mansfield, 1986).
- Attempted and Actual Assault Scale, ATTACKS (Bowers L, 2007; Bowers, Nijman, Palmstierna, & Crowhurst, 2002).

One of the disadvantages of “incident-based” scales is that they may not be easy to incorporate in ward routines (Sjöstrom, Eder, Malm, & Beskow, 2001), particularly when the target behaviour is rare. That is to say, the reliability of an incident-based registration method relies heavily on the preparedness of ward staff to record all aggressive incidents. “Period-based” aggression observation scales, intended for rating aggressive behaviour at predetermined times, may be less likely to be ‘forgotten’. Examples of period-based aggression observation tools include the Social Dysfunction and Aggression Scale, (SDAS) (Wistedt et al., 1990), and the
Ward Anger Rating Scale; WARS (Novaco, 1994). Such period-based aggression observation scales may, however, provide less information about the specific circumstances leading to incidents in the ward. For researching the effects of potential aggression-reducing interventions, a combination of ‘incident-based’ and ‘period-based’ aggression observation scales may be advisable.

The incident-based aggression observation scales such as MOAS, SOAS-R, REFA and ATTACKS all consist mainly of predefined answering options: aggression is recorded by ticking the options which apply to the observed behaviour. This makes them time-efficient as well as easy to use. The easier and quicker these scales are to complete, the more likely it will be that ward staff will record all incidents they encounter (ranging from mild to severe). In the following paragraphs, the MOAS, the SOAS-R, the REFA, and the ATTACKS instruments will be briefly described. In addition the RAGE, BARS and CMAI will be discussed with regard to their applicability in geriatric settings.

The Modified Overt Aggression Scale (MOAS)

The MOAS is an aggression observation instrument that divides aggression into four main categories, namely “verbal aggression”, “aggression against property”, “auto-aggression”, and “physical aggression” (Kay et al., 1988). Each of these four types of aggression is subdivided into five subcategories that reflect the severity of the behaviour (severity scores ranging from 0 to 4 points for each category). To calculate the overall severity of an incident, “aggression against property” scores are multiplied by a factor of 2, “auto-aggression” scores by a factor of 3, and
“physical aggression” scores by a factor of 4. They are then added to the “verbal aggression” score. In other words, the “physical aggression” score is given the greatest weight in calculating overall severity. In this way, total MOAS severity scores may theoretically range from 0 (no aggression) to 40 (most severe aggression). The inter-rater reliability of MOAS based on the total severity scores of two independent raters has been found to be good (Pearson’s $r = 0.85$ and 0.94; see Kay et al. 1988).

**Report Form for Aggressive Episodes (REFA)**

The REFA is a behavioural rating scale which measures aggressive behaviour toward others (Bjørkly, 1996). The REFA was specifically designed for the diagnostic purpose of mapping the situations that provoke aggressive behaviour in each patient. In other words, to an even greater extent than the SOAS-R, the focus of the REFA is on detecting (situational) triggers of aggression. The form consists of a list of 30 situations or interactions which might trigger aggression, grouped in the following seven main categories:

- Physical contact (four items),
- Limit setting (6 items),
- Problems of communication (3 items),
- Changes/readjustments (6 items),
- Persons (6 items),
- High-risk contact (2 items),
- Drugs/stimulants (3 items),
- An open category for additional situations or interactions.
Alongside the precipitants of aggression, there are six vertical divisions for the recording of defined characteristics of aggressive episodes: one for verbal threats, one for physical threats and four sections for physical assaults. As in the SOAS-R, the nurse who observed the event should record the aggressive incident on the form as soon as possible after the incident has taken place. However, for the REFA at least one other staff member must be consulted for a second opinion on the precipitants and the characteristics of the aggressive incident. After patients have calmed down, they are asked to provide information that may be of relevance for the accurate recording of the incident.

Two studies on the inter-rater reliability of the REFA both showed high inter-rater agreement and good reliability (Bjørkly, 1998). Fifteen years of clinical application indicate that the REFA has three clear advantages: accurate operational criteria for the definition of aggression, adequate emphasis on assessing situational variables, and recognition by many nurses as clinically useful. Finally, clinical experience indicates that it is easy to integrate REFA ratings in the planning and implementation of coping interventions and secure preventive measures in the clinical setting. The REFA also has several limitations: 1) it does not distinguish between serious and less serious physical assaults, 2) it does not ask whether the aggressive behaviour resulted in physical injury, 3) it does not ask what kind of measures were required to stop the aggressive behaviour, and 4) it has limited applicability to patients with very low rates of aggressive behaviour.
Rating Scale for Aggression in the Elderly: RAGE

The RAGE scale (Patel & Hope, 1992) measures the quantity and severity of aggressive behaviour, and involves a score on a four-point scale (0-3) for each of 21 items as well as a total score. Aggressive behaviour as shown by people with dementia is often brief and incidental, usually causing little harm to the victim, and for this group of patients it will often be practical to use an instrument that grades the prevalence of a certain behaviour over time. The RAGE is reported to be easy to use and focuses on observed behaviour. Its validity and reliability have been reported to be very good (Patel and Hope, 1992; Lam, Chui, & Ng, 1997; Shah, Chiu, & Ames, 1997). The authors of the scale define aggressive behaviour as: “an overt act, involving the delivery of noxious stimuli to (but not necessarily aimed at) another organism, object or self, which is clearly not accidental”.

The Cohen-Mansfield Agitation Inventory; CMAI

The Cohen-Mansfield Agitation Inventory (Cohen-Mansfield, 1986) is a 29-item caregiver rating questionnaire for the assessment of agitation in elderly persons. It includes descriptions of 29 agitated behaviours, each rated on a 7-point scale of frequency. Inter-rater agreement rates ranged between .88 and .92. A short version of this instrument, CMAI-SF, is also available. It contains 14 items to be rated by caregivers on a 5-point frequency scale. The items are based on the factor structure of the original inventory. This scale was designed for use in nursing homes, but it has also been used on acute psychogeriatric wards. The instrument was conceptualised as measuring agitation in two dimensions, verbal and physical, each of which has two poles, aggressive or non-aggressive. Studies based on this conceptualisation found
behavioural abnormalities of elderly subjects to consist of three main factors: verbally aggressive behaviour, verbally non-aggressive behaviour, and physically non-aggressive behaviour (Cohen-Mansfield, 1986, 1995). A fourth factor, physically aggressive behaviour, was added because of its importance in patient care, rather than because of the frequency of its occurrence (Cohen-Mansfield, 1995; Cohen-Mansfield, Werner, Watson, & Pasis, 1995).

The Attempted and Actual Assault Scale: Attacks

The Attacks scale (Bowers L, 2007; Bowers et al. 2002) is a more recently designed scale that seeks only to measure physical violence towards people, but it is intended to do so in great detail. Again, the Attacks scale should be completed by the staff members who have witnessed aggression, directly after a violent incident has taken place on their ward.

The central part of the scale is the most important and innovative feature of this instrument (Bowers, Nijman & Palmstierna, 2002). In this part, all physically violent actions that have been witnessed during an (attempted) assault are to be recorded. More specifically, all means and weapons used (e.g., sharp objects, hot liquids, spitting, poking) are to be noted in combination with the targets aimed at (e.g., head, limbs, torso, etc.). The frequencies of the separate physical actions are also to be estimated, as it can be assumed that striking more than once increases the likelihood of the violence causing serious harm.

One of the preliminary tests of the reliability and validity of the Attacks (Bowers, Nijman & Palmstierna, 2005) was done using video recordings of interpersonal assaults compiled from television broadcasts. During a meeting of the
European Violence in Psychiatry Research Group (EVIPRG), 22 members from 14 different countries were instructed to rate the videotaped assaults on both the MOAS and the Attacks. The inter-rater reliability of Attacks measurement of severity appeared to be promising.

**The Staff Observation Aggression Scale (SOAS)**

The SOAS (Palmstierna & Wistedt, 1987) is intended to measure verbal and physical aggression against objects, patients and/or staff. It was developed in Sweden to assess the degree and frequency of assaultive and violent behaviour in mental health settings. The SOAS comprises five columns pertaining to specific and consecutive aspects of aggressive behaviour (i.e., the provocation of the aggression, the means used during the aggression, the target of the aggression, the consequences of the behaviour, and the measures taken to stop aggression).

Every time a staff member has witnessed aggression by one of his or her patients, a SOAS form should be completed. The first column of the SOAS-R, which is intended to identify the factors which may have provoked or triggered the aggression, has been assumed to increase staff members’ sensitivity to risk factors in individual patients (see Nilsson, Palmstierna & Wistedt, 1988). Since 1999, a revised version of the SOAS (i.e., the SOAS-R; Nijman, Muris, Merckelbach, Palmstierna, Wistedt, Vos, van Rixtel, & Allertz, 1999) has been in use. This adapted version has a validated, more finely tuned system for scoring severity, which may increase the feasibility of comparing aggression rates between wards. In the original SOAS severity scoring system, a maximum score of 12 indicated the most severe incident. The revised SOAS-R severity score ranges from 0 to 22 points, higher scores again
indicating greater severity. The rationale for the revision of the severity scoring system was that the severity of aggressive behaviour depends on a range of factors, with some, such as the consequences for victims, being more important than others (such as the means used by the aggressive patient) in calculating the overall severity of incidents. Using regression techniques, a severity scoring system was developed in which separate features are weighted so that they make a differential contribution to the overall aggression severity score (Nijman et al. 1999).

Studies addressing the concurrent validity of SOAS and SOAS-R severity scores with other measures of aggression severity yielded significant results. For instance, the correlation with other methods for assessing the severity of aggressive behaviour varied from 0.38 to 0.81, (Nijman et al. 2005).

Cross-validations with clinical estimates of severity provided by staff members who had experienced aggression indicated that the revised SOAS-R severity scores approximate the general opinion in ward staff of the severity of incidents better than those in the original SOAS (Nijman et al., 1999; Nijman, Evers, Merckelbach, & Palmstierna, 2002)

Study results indicate fair to good inter-rater reliability for SOAS scores. On the basis of four incidents described, Palmstierna and Wistedt (1987) initially found an intra-class correlation of 0.96 between total SOAS scores from independent raters. Later studies conducted in clinical practice supported an acceptable inter-rater reliability for the scale, with Cohen’s $\kappa$s being 0.61 and 0.74 respectively (Nijman, Allertz, Merckelbach, á Campo & Ravelli 1997; Steinert, Wolfle, & Gebhardt, 2000) and a Pearson’s r between independent raters of 0.87 (Nijman et al., 1997). However, none of these studies examined the reliability of the decision when, or whether, to
complete a SOAS-R form between raters. As mentioned previously, the reliability of all incident-based aggression observation methods relies on the willingness of ward staff to complete a form after each aggressive occurrence. Particularly in cases of ‘mild’ aggressive behaviour, there may be inconsistency between raters in their decision on whether a SOAS-R form should be used to report the observed behaviour.

The SOAS measure was originally developed for use in geriatric settings, but soon its applicability was researched on acute wards. It has been undergoing a strict validation process and has been tested in different cultural settings. The form takes 2-3 minutes to fill in, and the findings can easily be used in evaluation processes at both the individual and institutional level, assisting the clinician in planning and developing care for patients with aggression-related problems and violent behaviour. SOAS or SOAS-R has also been used as the basis for a variety of incident forms nationwide, and its design is well known to the majority of clinicians in Norway. The simplicity of the form requires a minimum of training and education, which also makes it cost-effective and easy to implement. The use of SOAS-R to validate the BVC was a simple choice, despite the variety of observation-based scales available.
4.4 Research on inpatient violence in Norway

Until the mid-90s, few studies focused on violence, violence prediction and mental health in Norway. The main studies were the doctoral dissertations of Kirsten Rasmussen (1995) and Stål Bjørkly (1995), as well as an article on violence prediction by Linaker and Busch-Iversen (1995). Almvik and colleagues followed up the work by Linaker and Busch-Iversen (Almvik, Rasmussen, & Woods, 2006; Almvik & Woods, 1999, 2003; Almvik, Woods, & Rasmussen, 2000, 2007; Woods & Almvik, 2002).

During the past few years, an increasing number of studies on violent behaviour and related issues have been published in Norway, though few are directly related to violence within Norwegian psychiatry. This chapter will provide a brief synopsis of both major and recent works on inpatient violence in psychiatry in Norway.

In a study from a maximum secure unit, Rasmussen and colleagues (1995) compared 33 aggressive schizophrenic patients with 13 non-aggressive schizophrenic patients and 13 healthy controls. They used case history data, ratings of psychopathology, schizophrenic symptoms and neuropsychological tests. Findings revealed that the aggressive schizophrenic patients had spent more time in prison than the non-aggressive ones; they had earlier symptom onset and scored significantly higher on Psychopathy. They also displayed a specific pattern of neuropsychological dysfunction consistent with frontal lobe dysfunction, and their close relatives were more likely to have been involved in crime and substance abuse.
In another study of 94 consecutively admitted patients, Rasmussen and Levander (1996a) found that five types of crime and aggression could be identified through factor analysis: one non-violent type; one type with aggression/violence in an institutional setting; one with sexual violence; one homicidal; and one with arson as the major feature. Moreover, aggression within an institutional setting was the only type related to psychopathology. Psychopathy correlated strongly with both non-violent and violent crime, as well as with early adjustment problems. Arson was not explained by any of the background or symptom variables.

In a study of assaults on staff in a maximum-security psychiatric hospital, Rasmussen and Levander (1996b) analysed assaults on staff as registered by incident reports over a 6½-year period. During this time 94 patients were admitted to the unit. Fifty-two (55%) of the patients engaged in assaultive behaviour, generating 1945 incident reports. A small number of patients were responsible for a large number of the incidents. Serious incidents were rare. Incidents were evenly distributed throughout the day, week and year. Patients who attacked less often caused more serious harm and showed a decrease in assaults over time; the preceding events suggested that the violence was functional. For the more frequently assaultive patients it was difficult to identify a preceding event. The preceding events that were identified seemed more unreasonable, and the frequency of assault was constant over time. Patients who were more frequently assaultive were more often women, had more positive and borderline symptoms, were younger, and scored lower on psychopathy and depressive symptoms.

Bjørkly (1993) designed a Scale for the Prediction of Aggression and Dangerousness in Psychotic Patients (PAD). The scale is based on an interactional
understanding of aggressive behaviour in psychotic patients. The model emphasises
detailed analysis of each patient’s situational and interactional vulnerability, in
addition to personal variables, for improved prediction of aggressive behaviour.
Situational vulnerability in this context is defined as an increased likelihood to act
aggressively towards others in a given interaction. Based on 29 items grouped in 7
main categories, the PAD scores describe a patient’s profile of interactional
vulnerability.

Linaker and Busch-Iversen (1995) studied predictors of imminent violence in
psychiatric inpatients. They studied the behaviour and symptoms seen in the 24-hour
period preceding violent episodes. Six behaviours were more common before
violence: confusion, irritability, boisterousness, physical threats, verbal threats and
attacks on objects. A logistic regression equation based on these behaviours in a
randomised half of the observations predicted the occurrence of subsequent violence
in 92.1% of the other half of the sample without any false positives, giving a
sensitivity of 81.3% and a specificity of 100%.

Linaker (2000) also investigated the prevalence and characteristics of
dangerous female psychiatric patients. He conducted a national survey covering all
Norwegian inpatient and outpatient units for adults, including nursing homes and
prison services. A total of 329 people (male and female) were reported to have a
psychiatric disorder and to meet the survey’s criteria for dangerousness. Of these, 54
were females. When compared to a matched sample of men, the females had an
increased frequency of suicide- and self-injury-related behaviours as well as previous
episodes of arson.
Morken, Linaker and Langsrud (1998) asked what we can learn from variations in patient-staff incidents on acute psychiatric wards. From 1990, they recorded the time and day of the staff-patient incidents in order to examine time patterns and incident frequency. From a total of 653 incident reports, they selected the 347 which had, as a minimum, resulted in pain. They found differences in frequency of injuries during the year, with peaks in June, October and November. They also found differences during the week, with more injuries on weekdays than weekends. The nurses were more often interrupted in their work with patients early in the day on weekdays, which was the time at which most injuries occurred. There was a positive correlation between disturbances in the contact between staff and patients and the frequency of injuries. They concluded that the variation during the week and during the day was related to social rhythms. By analysing SOAS data and variations in the frequency of patient-staff incidents, the staff on a given ward can find out when the ward is functioning at its best. They also suggested that the variation during the year could depend on social or biological rhythms.

Urheim and VandenBos (2006) used the SOAS to study the pattern of aggressive and violent behaviour in a maximum-security hospital over a ten-year period, including 51 patients who generated a total of 4632 aggressive incident reports using the SOAS. Twenty percent of the patients caused 80% of the incidents and most of the incidents appeared in limit-setting situations. Female patients caused 50% more violent acts than their representation in the population in question, but they were less dangerous than the male population. Interestingly enough, only 36% of the incidents were reported to have no precipitants. The findings in this study were consistent with previous research in Norway.
Hartvig and colleagues (Hartvig, Alfarnes, Ostberg, Skjonberg, & Moger, 2006) did a preliminary study on a brief checklist for assessing violence risk among patients discharged from acute wards. During a one-year period, all discharged patients were scored at discharge with a 33-item risk assessment form (the Preliminary Scheme – PS) and during a one-year follow-up period episodes of aggression and violence were monitored when possible. As this was a preliminary study, no conclusions were drawn other than that it appeared to be possible to develop a brief screening instrument tailored for acute units.

In his doctoral thesis Vaaler (Vaaler, Morken, Flovig, Iversen, & Linaker, 2006) describes a study of possible predictive factors for threats and violent incidents in a Psychiatric Intensive Care Unit (PICU) based on evaluations performed at admission to the unit. These evaluations included both clinical interviews and standardised instruments such as the Global Assessment Functions Scale, Split version (GAF-S), Positive and Negative Syndrome Scale (PANSS). They also used the Brøset Violence Checklist as an outcome measure. Vaaler and his colleagues concluded that a clinical global judgment combined with a structured observer rating scale (BVC) was effective and more suitable than actuarial data in short-term prediction of aggression.

A recently published study by Langsrud and colleagues (Langsrud, Linaker, & Morken, 2007) explored the pattern of injuries to body parts in patient-staff incidents. Assaultive episodes were registered using the SOAS. There were no differences in gender and educational level among the victimised nursing staff, but nurses were most often injured in upper extremities, while doctors were more often injured in the head.
This is possibly “the first study to examine the relationship between injured body parts in patient-staff incidents and different aspects of the staff and the incident” (p.125). They also investigated the frequencies of incidents and found that the main provoking situation that preceded an incident was that the patient was denied something. Staff members were the target of the aggression in 95% of the incidents. In addition, they found a median of 18 on the severity score system in SOAS-R, which ranges from 0 to 22. This is higher than in all other studies using the SOAS-R severity score.

In conclusion, the number of studies on inpatient violence in Norway is increasing, but still insufficient. Except for the study by Langsrud et al., the findings in Norwegian studies are largely in line with international research. The lack of research is probably due mainly to lack of resources, but unlike many other countries, Norway has greater shortages of human resources (skilled professionals) than of funding.

By increasing its focus on research in this area, Norway – with its cultural and socioeconomic homogeneity (compared with many other countries) and presumably less variance due to societal factors – could probably add valuable knowledge.
5.0 Developing a tool for violence risk assessment

5.1 Developing the Brøset Violence Checklist

The Brøset Violence Checklist (BVC) was developed on the basis of the empirical work of Linaker and Busch-Iversen (1995). They investigated episodes involving physical violence documented in the medical records of 92 patients admitted to a Norwegian maximum-security unit during the years 1988-1993. The criterion for inclusion was violence severe enough to require the use of physical restraint. In all, 48 episodes of violent behaviour by 32 patients were identified; 16 of these patients were involved in two episodes each. All the daily nursing reports during the five-year period were also examined at specified control dates for reports of symptoms and behaviours. Single behaviours that were significantly more frequent before violent incidents were identified and used to build a logistic regression equation. The occurrence or non-occurrence of violence was used as a dependent variable, while the behaviours constituted the independent variables.

The six most frequent behaviours that occurred in the screening period prior to an incident were confusion, irritability, boisterousness, physical threats, verbal threats, and attacking objects. Entering these six behaviours into a logistic regression equation indicated that all of them were predictive of violence and seemed to constitute frequent warning behaviours (see table below).

An inter-rater reliability test was also performed, showing that all the six key behaviours correlated significantly between the raters with an average percentage agreement of 90.3 %
Table 1. Frequencies of key behaviours at control conditions and during the hours before violent episodes (Linaker & Busch-Iversen, 1995).

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Control before violence n=47</th>
<th>Pre-violent hours n=48</th>
<th>Control after violence n=46</th>
<th>$\chi^2$</th>
<th>$P_{\leq}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confusion</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>13.06</td>
<td>0.01</td>
</tr>
<tr>
<td>Irritability</td>
<td>2</td>
<td>10</td>
<td>1</td>
<td>11.85</td>
<td>0.01</td>
</tr>
<tr>
<td>Boisterousness</td>
<td>5</td>
<td>13</td>
<td>4</td>
<td>7.35</td>
<td>0.05</td>
</tr>
<tr>
<td>Physical threats</td>
<td>3</td>
<td>21</td>
<td>3</td>
<td>28.45</td>
<td>0.0001</td>
</tr>
<tr>
<td>Verbal threats</td>
<td>4</td>
<td>22</td>
<td>4</td>
<td>28.49</td>
<td>0.0001</td>
</tr>
<tr>
<td>Attacks on objects</td>
<td>2</td>
<td>18</td>
<td>3</td>
<td>24.02</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Only 14.6% of the violent incidents taking place during the study period were not preceded by any of these six behaviours, and more than half of the incidents were preceded by more than one.

Following up this early work, Almvik et al. have conducted several studies in various settings in Norway (Abderhalden et al., 2004; Almvik & Woods, 1999, 2003; Almvik et al., 2000, 2007; Woods & Almvik, 2002), developing a more user-friendly interface for the BVC form as well as introducing a learning package and a manual describing previous research and providing a detailed description of each item and how to use the instrument.

The BVC is based on the idea that violent incidents rarely occur without identifiable warning signals. It is one of the few validated structured clinical instruments developed for an inpatient setting and one of very few instruments that assist in here-and-now risk assessment during inpatient treatment.
5.2 *International validation of the BVC*

The first two works published on the development of the BVC attracted considerable international attention, and numerous collaborative works have been produced.

In Australia, one attempt to develop risk assessment practices for inpatient aggression, Ogloff and Daffern (2006) included items derived from antecedent interactions and purposes (‘sensitivity to perceived provocation’, ‘unwillingness to follow direction’, and ‘easily angered when requests are denied’) and compared these with the Brøset Violence Checklist and items taken from the clinical and risk scales of the HCR-20.

Results of this study showed that the Area Under Curve (AUC) for the BVC was 0.83, while the AUC for the C-scale in HCR-20 was 0.73. Three items derived from the study performed as well as or better than many of the other items from the HCR-20 and Brøset Violence Checklist, with AUC statistics of .76 for ‘unwillingness to follow direction’, .75 for ‘sensitivity to perceived provocation’, and .74 for ‘easily angered when requests are denied’. Only irritability (.77) and impulsivity (.77) had better AUCs. As a result of this research, these three items were integrated into the Dynamic Appraisal of Inpatient Aggression, a risk assessment system developed to assist in the assessment of imminent aggression. From this work Ogloff and Daffern developed the Dynamic Appraisal of Situational Aggression (DASA), but the AUC values for the DASA and the BVC were not significantly different from each other.

Abderhalden in Switzerland (Abderhalden et al., 2004) conducted a study involving six acute wards in three psychiatric hospitals in Switzerland using the German version of the BVC (BVC-G). The total sample consisted of 219 patients...
consecutively admitted to the units (61% male, 39% female, mean age 39.9 years). Scores were reported to be skewed towards the lower end of the BVC-G total. Using a cut-off point of a score of 3 or more on the BVC-G and the likelihood that a violent incident would occur in the following 12-hour period, they reported 6.4% false predictions, sensitivity 64%, specificity 94% and an AUC of 0.88. Abderhalden and colleagues also studied inter-rater reliability, where four nurses independently rated 16 patients. Reliability ranged from $r = 0.64$ to $r = 1.0$ at the item level. In a study published in 2006, Abderhalden and his group aimed at improving the accuracy of short-term violence prediction by combining the BVC with an overall subjective clinical risk assessment and testing the application of the combined measure (Abderhalden et al., 2006). The subjective measure consisted of a 6-point score on a Visual Analogue Scale (VAS). Although this study showed that the modified BVC was an accurate instrument in short-term prediction of violence in acute psychiatric care, the inclusion of the VAS-derived data did not change the accuracy of the original BVC.

In Sweden, Palmstierna and Olsson (2007) did a study on violence from young women involuntarily admitted for severe drug abuse at 3 specialised treatment institutions in Stockholm. Their aim was to examine the predictive capacity of both static and dynamic risk factors for imminent violence (24-hour perspective) as well as to investigate the effect of different institutional settings. In this article they also introduced an extended Cox regression model allowing them to look at multiple events from the same patients as well as time-dependent covariates. In this material, the Cox model revealed that the most important factors for short-term aggression prediction were all dynamic factors. However, except for the item “attacking objects”,
none of the other items could further explain or predict severe violence within the
next 24 hours. The authors suggest that the BVC has the capability to predict
imminent violence, but that its properties vary in different settings and ward structures
and that the results in this study could be due to differences in design or statistical
analyses.

Using the extended Cox model, Bjørkdahl and colleagues (Bjorkdahl, Olsson, &
Palmstierna, 2006) evaluated the short-term capacity of the BVC when used by nurses
at a 10-bed psychiatric intensive care unit in Sweden. Eleven (15.1%) out of 73
patients admitted during a 3-month period were reported to be severely violent
according to the SOAS-R severity measure. A total of 997 daily predictions with the
BVC were reported. Of these, 758 BVC ratings gave a total score each of zero, 239
scorings with a sum of at least 1. Their analysis showed that any positive score on any
item of the BVC resulted in a six-fold increase in the risk of severe violence, while a
negative score on all BVC items correctly predicted no risk for severe violence in
99.2% of all assessments. Björkdahl et al. concluded that “short-term risk for severe
inpatient violence was to a high degree accurately predicted by nurses using the
Bröset Violence Checklist (BVC)”.

In summary, the studies from Australia, Switzerland and Sweden mentioned
above provide strong evidence of very good predictive validity of the BVC,
confirming the results from early and recent Norwegian studies.
6.0 Methods

6.1 Methodological considerations

Data collected in the present studies (paper 1, 2, 4 and 5) are based on staff observations. This is considered valid both for BVC data and for the data on violent incidents collected using SOAS-R. One could argue that this is limited to the extent that staff observing either a violent incident or a different kind of behaviour (e.g. the six BVC items) are able to provide an accurate account of it. However, studies conducted on both scales consistently show fair to good inter-rater reliability (Almvik, Woods, & Rasmussen, 2000; Nijman, Palmstierna, Almvik, & Stolker, 2005).

In all the studies, staff were instructed to act as usual in their clinical practice. The staff members were trained to score the BVC and the SOAS-R, but they were not aware of the interpretations of the BVC scores. Furthermore, all BVC observations were scored prior to any incident and were hence inherently prospective. With the exception of data collection, treatment took place as usual. Informal interviews with the staff after the pilot study confirmed that this was the case.

Missing data could be a concern. Missing values and subjects seem to be common in studies conducted in naturalistic settings. We could not, however, identify any systematic bias in missing data in relation to sites or shifts. Although a possible bias cannot be ruled out, the prospective nature of the study, as well as the considerable number of reports and consistency in pattern across populations, would argue in favour of the representativeness of the data.
6.2 Statistics

In practice, risk can always be described in dichotomous terms: the patient will or will not be violent. A common way of defining precision, which was developed more than 50 years ago in radiology (Rothwell et al. 2005), applies the terms sensitivity and specificity. Usually research on the prediction of violence has been reported using two-by-two contingency tables to display results. This then allows for the examination of correct predictions and error rates. Two possible outcomes exist in this design: either violence did occur or it did not. Typically this is reported as a **true positive** (the patient was predicted to be violent and was violent), and a **false positive** [type I error] (the patient was predicted to be violent and was not violent), a **true negative** (the patient was predicted as non-violent and s/he was not violent) or a **false negative** [type II error] (the patient was predicted to be non-violent and was violent).

The sensitivity of an instrument (TP/ [TP+FN]) is concerned with the correct prediction of the outcome occurring, i.e. violent behaviour. In contrast, the specificity (TN/[TN+FP]) is concerned with the correct prediction of the outcome not occurring.

Table 2. Prediction outcome

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Violent</th>
<th>Not violent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violent</td>
<td>True Positive (TP)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>Not violent</td>
<td>False Negative (FN)</td>
<td>True Negative (TN)</td>
</tr>
</tbody>
</table>
The terms Positive Predictive Value (PPV) and Negative Predictive Value (NPV) are also used. In the model above, PPV can be described as $\frac{TP}{TP+FP}$ and the NPV as $\frac{TN}{TN+FN}$. Positive and negative predictive values are crucial in the evaluation of different kinds of models for risk assessment. They immediately provide a clear picture of “rights” and “wrongs” in a clinical adjusted situation. However, a problem with the PPV/NPV approach is the sensitivity for the base rate of the phenomenon. Since violent incidents in psychiatry are rare, PPV is seldom high.

In the acute ward study (see paper I) the figures illustrated as a 2x2 table would look like this with outcomes for a score of 2 and above on BVC ratings ($n=901$):

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Violent</th>
<th>Not violent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violent</td>
<td>40</td>
<td>66</td>
</tr>
<tr>
<td>Not violent</td>
<td>24</td>
<td>771</td>
</tr>
</tbody>
</table>

This would yield a PPV of 0.38 and a NPV of 0.97

In 1994 Mossman (Mossman & Somoza, 1991) suggested a new statistical approach to risk assessment: Receiver Operating Characteristic analysis (ROC analysis). The method had been developed during the second world war to describe the relationship between signal and noise in radar technology and had for a long time been used in epidemiological research. Mossman suggested modifying the analysis, adjusting it to research on prediction of violence among psychiatric patients. ROC analysis is based on plots on a curve where each plot describes the balance that appears in a diagnostic instrument between sensitivity and specificity for all possible threshold values. The key value for interpreting a ROC analysis is the area under the
curve (AUC). The better the classification, the further the curve is from the straight diagonal line – the “by chance” alternative. ROC analysis allows for assessments at several sensitivities and specificities, not just the one pair. As used in figure 1, plots can be shown visually and also compared visually by adding several plots in one curve.

Figure 1: example of a ROC curve

Although ROC analysis is seen more and more in scientific papers, use of this method still entails some problems. One problem that has been pointed out by Sjøstedt and Grann is that there is no strong consensus on interpreting AUC estimates for predictive validity and that there seems to be a tendency to interpret the AUC too optimistically. Given the possible human and monetary costs associated with errors,
they therefore suggested that AUCs should be interpreted conservatively: AUC < 0.60 = low precision; AUC 0.60 - 0.70 = marginal precision; AUC 0.70 - 0.80 = modest precision; AUC 0.80 - 0.90 = good precision and AUC > .90 = high precision (Sjöstedt & Grann, 2002). A more in-depth analysis of current BVC studies indicates that although the precision of the instrument may vary depending on the interval from observing and scoring to the occurrence of the event (i.e. the violent incident), results so far show moderate to good precision and even high precision on BVC scores recorded on the same shift as the incident.

### 6.2.1 Summary

To assess or estimate risk is a very difficult task, but success will benefit staff, patients and society. Staff will feel, and actually be, safer; patients will benefit from improved clinical environments; and the general anxiety in communities for people with a mental illness will be reduced. There is no clear pathway to best practice in this area, but improvements in research quantity and quality will obviously bring us one step closer to an even better way of reducing violence in psychiatry. Also, ongoing debate in both professional and open public forums is needed to keep clinicians, researchers and the public informed about pros and cons and updated on the latest development in the field of risk assessment.
7.0 Summary of papers

7.1 Paper I

The Brøset Violence Checklist (BVC) and the prediction of inpatient violence: some preliminary results.

The objective of this study was to test the BVC in acute wards and to report early results and psychometric properties for the checklist. During a two-month period, all patients who were admitted to three acute wards in three different psychiatric hospitals in Norway were rated using the BVC. The patients were rated on the day of admission and on each shift of the following three days. Data on incidents were collected using the reporting system utilised by all wards, which is based on the Staff Observation Aggression Scale. During the study period 109 patients (52 males and 57 females) were admitted and rated on the participating wards; Twelve patients were reported to have been involved in an incident over the period of study; four (33.3%) were male and eight female (66.7%). These patients were involved in a total of 32 violent incidents.

Results showed that a score of two and above was predictive of a violent event. Of all the ratings, 136 would predict that violence would occur in the next 24 hours, and 734 that it would not. The sensitivity was found to be 0.74 or 74% accurate at detecting the outcome of a violent event for a score of two or above.
For the BVC, the specificity was found to be 0.91 or 91% accurate at detecting the outcome of a violent event not occurring for a score of 1 or below.
7.2 Paper II


This paper explored the inter-rater reliability, sensitivity and specificity of the BVC in acute wards. All 109 consecutive referrals to three psychiatric inpatient acute units during a two-month period were included in the study. Ratings were performed at the time of admission and three times a day for each patient – once for each working shift. Thirty-four separate incidents of violence occurred.
Two methods were used to examine inter-rater reliability on a set of ratings by two independent raters (N=39), Kappa values and percent exact. The Kappa values for the six BVC items varied from .48 to 1.00 while the Kappa value for the total BVC score was .48. A second method used was percentage exact rater agreement, which is a complementary method for categorical data. Findings ranged from 90 to 100% agreement on the six items.
Inter-rater reliability was adequate to good; comparisons between ratings performed in the 24-hour interval before the incident, and all other ratings suggested moderate sensitivity and good specificity of the instrument. It was concluded that the BVC was a useful instrument in predicting violence within the next 24-hour period and that the psychometric properties of the instrument are satisfactory.
7.3  **Paper III**

Fifteen years of research with the Staff Observation Aggression Scale; A review
111: 12–21

A literature search was conducted with the aim of collecting all papers
between 1987 and 2001 in which SOAS or SOAS-R data were published. From
this material, all findings about the psychometric properties of the SOAS and
SOAS-R were extracted, as well as the average number of incidents per patient
per year, and the mean SOAS or SOAS-R severity scores. The aggression
frequencies and severity scores were summarised and compared.
The review of psychometric studies indicates satisfactory inter-rater reliability and
validity of SOAS(-R) assessments. The annual frequency of aggressive incidents
caused by adult psychiatric patients residing in acute admissions wards varies
across studies, from 4.4 (Italy) to 33.0 (Netherlands). Dutch frequencies may in
general be somewhat higher than those found in other countries. Interestingly, the
highest prevalence rates (i.e. 57.6 incidents per patient per year) were reported for
psychogeriatric wards in Sweden. The mean average severity of incidents in terms
of SOAS(-R) severity scores is highly stable across countries.

Conclusions: The stability of the mean severity reported for incidents across
countries and ward types suggests that ward team members from different
countries apply fairly similar criteria in deciding which behaviour should be
reported. With the same mean severity, however, prevalence rates across countries
varied considerably. Such differences may be the result of different national
policies on what treatment is allowed to prevent dangerous situations (e.g., forced medication or not), but may also be the consequence of local differences in environmental ward variables (e.g., different staffing levels, private bedrooms or not). Although the preliminary data collection does not allow firm conclusions, the uniform use of SOAS(-R) across cultures may provide new options to investigating how local situational variables are associated with aggression frequencies.
7.4 **Paper IV**

Challenging behaviour in the elderly - monitoring violent incidents.


The aim of the study was to explore the frequency and nature of violent incidents in psychogeriatric wards and nursing homes in terms of type and severity of incidents, what provoked the incidents, and what kind of measure was needed to stop the aggression. Aggressive behaviour of the study group was monitored using the Staff Observation Aggression Scale-Revised (SOAS-R) in two Norwegian nursing homes and two geriatric psychiatric wards for a period of three months. Severity of incidents was monitored with the built-in severity scoring system in SOAS-R. During the study period, 32 out of the 82 patients were reported to be violent. A minority of the patients generated the majority of the incidents. Physical injury to the staff as a consequence of the aggression was extremely rare. Situations where the client or patient was denied something were the most provocative, and a substantial number of incidents occurred at bath/shower times. Talking to the patient was the most frequent measure used to stop the aggression, but more intrusive measures were also used. Conclusions: A substantial proportion of the incidents was associated with personal care tasks, suggesting a crucial role for communication difficulties and a focus for staff training. We suggest that personal care situations should be added to the variable list in future research.
7.5 **Paper V**

Assessing risk for imminent violence in the elderly; the Bröset Violence Checklist


This paper aimed to test the validity of the BVC in psychogeriatric and nursing home settings and to report on the predictive value of the instrument. In total, 8835 BVC observations were completed in two psychogeriatric wards (n=42 patients) and two special-care units for patients with dementia (n=40 residents). To measure violent incidents, the subjects were monitored using the Staff Observation Aggression Scale-Revised (SOAS-R). This study showed that patients in geriatric wards and residents in nursing homes who were aggressive had higher BVC scores than the non-violent subjects, indicating that the BVC does predict violent episodes in these settings.

Conclusion: From a clinical perspective, it is most important that a prediction aid has good sensitivity, so that most cases are detected, and that it has a high negative predictive value so that most non-cases on the measure are indeed non-cases. Our results indicate that the BVC was able to achieve this goal.
8.0 Discussion

8.1 Clinical implications and future research

Within psychiatric inpatient units, vast time commitments to regular monitoring of potentially violent behaviour are not feasible. It thus appears that there is a need for a quick and easy-to-use instrument to predict the risk of imminent violence. One of the fundamental advantages of using the BVC in everyday clinical practice is that it takes only a couple of minutes to complete, and easy-to-use instruments are obviously more likely to actually be used on an everyday basis in a busy clinical ward environment.

Violence prediction and risk assessment have been criticised for a number of reasons, but for the most part critics have been concerned about the consequences of false positive predictions, i.e. when a patient has been predicted to be violent and no violence occurred. The consequences of such predictions will vary according to the setting and purpose, with the extreme represented by Dr. Death as described by Hare (Hare, 1993), where prisoners could be sent to death row on the basis of risk assessment by the Psychopathy Checklist. Less fatal, though still unwanted, consequences are when patients are denied discharge or prisoners detained on the basis of what might be false positive predictions.

In the case of assessing risk of violent behaviour among inpatients in the immediate future by the BVC, the consequences of possible false positives should be less intrusive. Interventions reported in the present study are typically one-to-one observation, seclusion, extra staff called in, or additional or changed medication.
Although some would argue that seclusion and medication are intrusive measures, most people would say they are inferior to the measures applied by the legal system.

The BVC is designed to detect and prevent violence in health care settings in a way that is useful for clinicians, administrators and service users/patients. It is very brief and easy to use; in practice there is no burden upon service users/patients. Due to its simplicity, training also takes a matter of minutes rather than hours or days, and hence the BVC is also cost effective. Stedman and colleagues (Stedman et al., 2000) have proposed criteria or dimensions for assessing the suitability of an instrument:

1. **The measure must be applicable.** Both the research referred to in this thesis and clinical feedback clearly indicate that the instrument is seen as applicable by mental health care staff.

2. **The measure must be acceptable.** It should be brief, and the purpose, wording, and interpretation should be clear. As mentioned above, the BVC is indeed brief; both the wording and interpretation are clear and short.

3. **The measure must be practical.** Issues of practicality relate to the burden imposed on consumers and service providers in terms of time, costs, training and level of skill required in the scoring and interpretation of the data. Compared to most risk assessment scales available, the BVC requires a minimum of training and is multidisciplinary in its nature.

4. **The measure must be valid.** The BVC shows sound psychometric properties and measures what it is supposed to measure. International studies also show that the BVC can be used in various cultural settings with results just as good as in the Norwegian research.
5. The measure must be reliable. Reliability data indicates fair to good inter-rater reliability (Almvik, Woods, & Rasmussen, 2000). Both research and clinical experience shows that the BVC provides the same results when given to two different people rating the same person at the same time.

As the ultimate goal of risk assessment is violence reduction, future research should investigate whether the introduction of BVC in a psychiatric ward actually results in a reduction in violent episodes. Furthermore, a more systematic approach to “what works for whom” should be undertaken; i.e. do different measures work differently in patients who score high on irritability versus patients who score high on other items? Should psychiatric units have fixed standards for procedures given a certain BVC score, and will that reduce violence? Also, increased research activity will render a basis for the development of a national and international strategy for predicting, reducing and handling violent incidents.

Currently a software version of the BVC is under development and has been trialled in forensic settings in Denmark. The introduction of an electronic BVC could further increase the usability of the interface. Access to BVC data, for example from previous admissions, could also improve the clinical applicability of the instrument. As the electronic version is still at a trial stage, future research should explore any advantages or disadvantages of such an implementation.

As listed above, present research clearly indicates that the BVC is a quick and easy-to-use aid in informing clinicians about imminent violence in various clinical setting. Extensive international research indicates good validity and reliability, and feedback from clinicians worldwide is very good. The next step should include an
investigation into its applicability to the goal of risk reduction as well as tests of the BVC in an even wider range of clinical settings.
References


European Foundation for the Improvement of Living and Working (2005)


Errata

Please note that in papers I and II the specificity and sensitivity results reported for a score of 2 or more as being predictive of violence are different, even though they come from the same data set. The reason for this is that for Paper II it was decided to employ more stringent criteria when recoding for an incident occurring during the next 24-hour period.

Please also note that in Paper II there is a typographical error on page 1288 and unfortunately none of the authors became aware of this until recently. However, this does not change or in any way affect the results and conclusions of the paper. The correction is as follows:

Text in Paper II:

“Consequently, this left 64 ratings where there was a score of 2 or above and an incident occurred in the next 24 hours. In summary, the behavior of confusion showed in 50% of the ratings, irritability in 58%, boisterousness in 53%, verbal threats in 34%, physical threats in 23%, and attacks on objects in 33%. As for the 837 ratings below a score of 2 in which no violence occurred, confusion showed in 16% of the ratings, irritability in 7%, boisterousness in 5%, verbal threats in 2%, physical threats in less than 1%, and attacks on objects in just above 1%.”
Corrected text:

“Consequently, this left 64 ratings where an incident occurred in the next 24 hours. In summary, the behavior of confusion showed in 50% of the ratings, irritability in 58%, boisterousness in 53%, verbal threats in 34%, physical threats in 23%, and attacks on objects in 33%. As for the 837 ratings in which no incident occurred, confusion showed in 16% of the ratings, irritability in 7%, boisterousness in 5%, verbal threats in 2%, physical threats in less than 1%, and attacks on objects in just above 1%.”
Papers are not included due to copyright.
Appendix
**The Brøset Violence Checklist © (BVC) - quick instructions:**

Score the patient at agreed time on every shift. Absence of behaviour gives a score of 0. Presence of behaviour gives a score of 1. Maximum score (SUM) is 6. If behaviour is normal for a well known client, only an increase in behaviour scores 1, e.g. if a well know client normally is confused (has been so for a long time) this will give a score of 0. If an increase in confusion is observed this gives a score of 1.

<table>
<thead>
<tr>
<th>Monday</th>
<th>Tuesday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confused</td>
<td>Confused</td>
</tr>
<tr>
<td>Irritable</td>
<td>Irritable</td>
</tr>
<tr>
<td>Boisterous</td>
<td>Boisterous</td>
</tr>
<tr>
<td>Verbal threats</td>
<td>Verbal threats</td>
</tr>
<tr>
<td>Physical threats</td>
<td>Physical threats</td>
</tr>
<tr>
<td>Attacking objects</td>
<td>Attacking objects</td>
</tr>
<tr>
<td>SUM</td>
<td>SUM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wednesday</th>
<th>Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confused</td>
<td>Confused</td>
</tr>
<tr>
<td>Irritable</td>
<td>Irritable</td>
</tr>
<tr>
<td>Boisterous</td>
<td>Boisterous</td>
</tr>
<tr>
<td>Verbal threats</td>
<td>Verbal threats</td>
</tr>
<tr>
<td>Physical threats</td>
<td>Physical threats</td>
</tr>
<tr>
<td>Attacking objects</td>
<td>Attacking objects</td>
</tr>
<tr>
<td>SUM</td>
<td>SUM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confused</td>
<td>Confused</td>
</tr>
<tr>
<td>Irritable</td>
<td>Irritable</td>
</tr>
<tr>
<td>Boisterous</td>
<td>Boisterous</td>
</tr>
<tr>
<td>Verbal threats</td>
<td>Verbal threats</td>
</tr>
<tr>
<td>Physical threats</td>
<td>Physical threats</td>
</tr>
<tr>
<td>Attacking objects</td>
<td>Attacking objects</td>
</tr>
<tr>
<td>SUM</td>
<td>SUM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confused</td>
</tr>
<tr>
<td>Irritable</td>
</tr>
<tr>
<td>Boisterous</td>
</tr>
<tr>
<td>Verbal threats</td>
</tr>
<tr>
<td>Physical threats</td>
</tr>
<tr>
<td>Attacking objects</td>
</tr>
<tr>
<td>SUM</td>
</tr>
</tbody>
</table>
The Brøset Violence Checklist

Interpretation and Operationalisation

Interpretation of scoring:

Score = 0  The risk of violence is small
Score = 1-2  The risk of violence is moderate. Preventive measures should be taken.
Score > 2  The risk of violence is very high. Preventive measures should be taken. In addition, a plan should be developed to manage the potential violence.

Operationalisation of behaviours/items:

<table>
<thead>
<tr>
<th>Confused</th>
<th>Appears obviously confused and disorientated. May be unaware of time, place or person.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irritable</td>
<td>Easily annoyed or angered. Unable to tolerate the presence of others.</td>
</tr>
<tr>
<td>Boisterous</td>
<td>Behaviour is overtly &quot;loud&quot; or noisy. For example slams doors, shouts out when talking etc.</td>
</tr>
<tr>
<td>Physically threatening</td>
<td>Where there is a definite intent to physically threaten another person. For example the taking of an aggressive stance; the grabbing of another persons clothing; the raising of an arm, leg, making of a fist or modelling of a head-butt directed at another.</td>
</tr>
<tr>
<td>Verbally threatening</td>
<td>A verbal outburst which is more than just a raised voice; and where there is a definite intent to intimidate or threaten another person. For example verbal attacks, abuse, name-calling, verbally neutral comments uttered in a snarling aggressive manner.</td>
</tr>
<tr>
<td>Attacking objects</td>
<td>An attack directed at an object and not an individual. For example the indiscriminate throwing of an object; banging or smashing windows; kicking, banging or head-butting an object; or the smashing of furniture.</td>
</tr>
</tbody>
</table>

NB: For the behaviours/items physically threatening, verbally threatening and attacking objects the operationalisation was adapted from the Behavioural Status Index (Reed, Woods & Robinson, 2000) by one of the authors (Woods).

© Roger Almvik and Phil Woods 2000 – not to be copied without express written permission by one of the authors (email: roger.almvik@ntnu.no or phil.woods@usask.ca)