PREDICTION OF RISK OF HARM OF INPATIENT AGGRESSIVE BEHAVIOURS IN ST. JOSEPH'S HEALTHCARE HAMILTON

PREDICTION OF RISK OF HARM OF INPATIENT AGGRESSIVE BEHAVIOURS IN ST. JOSEPH'S HEALTHCARE HAMILTON

By

EMANUELE BLASIOLI, B.Sc., G.Dip., M.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2018)

Hamilton, Ontario eHealth Program

TITLE: Prediction of Risk of Harm of Inpatient Aggressive Behaviours in St Joseph's Healthcare Hamilton

AUTHOR: Emanuele Blasioli, B. Sc., G. Dip., M.Sc.

SUPERVISOR: Dr Elkafi Hassini

CO-SUPERVISOR: Dr Peter J. Bieling

NUMBER OF PAGES: xiv; 155

LAY ABSTRACT

Violence in inpatient psychiatric represents an issue for patients and clinical staff, and impacts significantly the quality of the health care services delivered. This retrospective study focuses on violence risk prediction at St. Joseph's Healthcare Hamilton, between 2016 and 2017. We studied the predictive performance of a clinical indicator, the Risk of Harm to Others Clinical Assessment Protocol (RHO CAP), embedded in a mandatory assessment tool for psychiatric facilities in Ontario, the Resident Assessment Instrument for Mental Health (RAI-MH). After analysing the performance of this indicator and studying the most important factors associated with harmful aggressions, we develop an algorithm to predict the risk of harm in inpatient settings. Our results stress the importance of the RHO CAP in support of a safer environment and a higher quality of care. The proposed predictive model is an attempt towards a further improvement for predicting inpatient risk of harm.

ABSTRACT

Inpatient violence risk prediction is a priority for safety and quality control purposes. The aims of this retrospective study is: 1) analysing the performance of the Risk of Harm to Others Clinical Assessment Protocol (RHO CAP) in predicting the risk of harm within St Joseph's Healthcare Hamilton; 2) identifying the most important risk factors associated with harmful incidents in inpatient mental health; 3) developing an alternative algorithm to predict the risk of harm in inpatient settings; 4) analysing the performance of the RHO CAP among patients who did not commit any aggression. Data from January 2016 to December 2017 have been anonymized and collected, for a total of 870 episodes of inpatient aggressions perpetrated by 337 patients. Two main sources of information have been used: the Resident Assessments Instrument for Mental Health (RAI-MH), and an internal report of the aggression incidents. We develop a Bayesian probabilistic classifier, a logistic regression, to investigate the risk factors for harmful incidents and propose a predictive model for risk of harm.

The RHO CAP has demonstrated a better performance in discriminating which patients were more at risk to commit some type of aggression than at identifying the risk of harm among those who will commit aggression. The factors most significantly associated with harmful incidents were *age, history of violence to others, police intervention for violent behaviour, and a diagnosis of psychosis*. The proposed predictive model showed an overall accuracy of 75%. It focuses on four diagnoses (the most frequent diagnoses associated to aggression in our sample), age, history of violence to

others and police intervention for violent behaviour, and inappropriate behaviour within the social context.

ACKNOWLEDGMENTS

I would like to acknowledge Dr Elkafi Hassini's valuable supervision during the whole program of study, all his aid and scientific support made possible the realization of the entire pathway at McMaster University.

I would also like to acknowledge my co-supervisor Dr Peter J. Bieling, whose insights and support have been critically important for the realization of this study.

I would like to acknowledge the Hamilton Integrated Research Ethics Board (HIREB) for having approved this study.

A special thanks to the committee chair, Dr Norm Archer, for his comments and recommendations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DECLARATION OF ACADEMIC ACHIEVEMENT

I, Emanuele Blasioli, declare this thesis to be my own work.

I have received support and guidance by my supervisors, Dr Elkafi Hassini and Dr Peter J. Bieling, as well as by Dr Norm Archer who reviewed my work providing meaningful comments and recommendations.

I declare, to the best of my knowledge, that this work does not infringe anyone's copyright.

Chapter I

1. Introduction

The topic of this thesis is violence in inpatient psychiatric settings, a complex phenomenon strictly connected with the risk of harm to which both patients and clinicians are exposed. In this context, the possibility to predict the risk of exposure to harm became more and more a priority, and it led to the development and implementation of various measures and tools that are utilized for many purposes, from patients' assessment to care planning.

As part of their continuous improvement efforts, St. Joseph's Healthcare Hamilton sought to review their current assessment tools for the risk of harm from aggression in inpatient psychiatric clinics. The idea of this thesis grew out of the author's internship work as part of a collaboration between the eHealth program at McMaster University and the St. Joseph's Healthcare Hamilton (SHH), the support of which has been critical for the successful completion of this research. SHH wanted to evaluate the predictability power of the Resident Assessment Instrument for Mental Health (RAI-MH), with specific focus on the algorithm for risk of harm prediction, the Risk of Harm to Others Clinical Assessment Protocol (RHO CAP). Part of this study we have also identified the most important factors involved in harmful episodes of aggression and proposed an alternative predictive model. We have also looked the ability of the RHO CAP algorithm in discriminating residents who represented a minimal or null risk to commit aggressions among those who did not commit any aggression.

In this thesis research we have made several contributions: (1) We provide a real word case study to evaluate the performance and behaviour of the RHO CAP, as well as other clinical factors. To our knowledge, this is the first study that investigates the predictors of harm using the RAI-MH. (2) We develop a predictive model for risk of harm in patients who have committed aggression. This has shed light on what directions can be pursued in the future to bring improvements into the field of violence prediction. (3) Evaluate the performance of RHO CAP for predicting aggression for patients who did not commit prior aggression. Our work can aid clinics in assessing the probability of aggression occurrence as well as its time.

The remainder of this thesis is organized as follows. In the second chapter we review the relevant literature on aggression and violence in mental health. We define aggression and violence and discuss how they impact the mental health care sector. The chapter continues with a section that illustrates the complexity of violence and aggression within the psychiatric setting, the predictors for the inpatient risk of harm that emerged from previous research. We conclude the chapter with a discussion on the management of violence and aggression in psychiatric settings.

In the third chapter we discuss four important psychological theories of violence and aggression, starting from the one developed by Freud in the beginning of the twentieth century to the most recent theory of aggression, the General Aggression Model (GAM; Anderson and Bushman, 2002), which is considered as the most comprehensive and integrative framework. The psychological theories are followed by a summary of the main structural and functional neuroanatomical findings related to aggression and violence in human beings.

In the fourth chapter we discuss the future challenges for the healthcare systems and the main reasons for a transition towards new patient-centric models of healthcare. We then provide an overview of history, development and adoption of the RAI instruments. This is further detailed in chapter five, with a focus on the Residential Assessment Instrument for Mental Health (RAI-MH), which we use for our retrospective study.

In the sixth chapter we introduce the concept of violence risk assessment and prediction, and present the main statistical methods adopted in violence risk prediction: the logistic regression, classification trees and random forests, and neural networks models.

In the seventh chapter we make use of our findings from the literature in the previous chapters as well as data and context from the St Joseph's Healthcare Hamilton to evaluate the performance of RAI-MH in predicting the risk of harm in patients who committed aggression and develop a predictive model for risk of harm for such patients. We also assess the prediction power of RAI-MH in classifying aggression behaviour for patients who did not commit prior aggression and develop a probability model that determines the chances that a patient will commit aggression within a given period of time.

We provide our conclusions, study limitation and future research directions in chapter 7.

Chapter II

Aggression and violence in mental health care

1. Background

Violence affects significantly the quality of the health care services delivered. The risk of being exposed to violence is of concern for both employees and patients in an inpatient psychiatric setting (Quintal, 2002). Despite the many reforms in health care and deinstitutionalization initiatives during the past decades, violence in psychiatric wards has increased. Different studies show that aggressive and violent behaviours continue to rise all over the world, affecting all healthcare professionals (Cutcliffe & Riahi, 2013a).

The rates of violence in inpatient settings reported in the literature vary significantly. Despite these variations, there is no doubt that the percentage of nonfatal assaults in psychiatric facilities is significantly higher than in all other industries, and the nursing staff is the personnel that is typically more affected, since they spend more time with patients than any other health care professional (McDermott & Holoyda, 2014). According to the Department of Justice National Crime Victimization Survey (US), which looked at nonfatal violence in workplace in the period 1993 to 1999 (Duhart, 2001), employees in mental health are the second most exposed category to nonfatal violent behaviours, with a rate of 68.2 violent victimizations per 1000 workers. More specifically, the report explains that:

> *"The workplace violent crime victimization rate for nurses was not significantly different from that for physicians; however, nurses experienced workplace crime at a rate 72% higher than medical technicians and at more than twice the rate of other medical field workers (22 versus 13 and 9, respectively). Professional (social workers/ psychiatrists) and custodial care providers in the mental health care field were victimized while working or on duty at similar rates (68 and 69 per 1,000, respectively) — but at rates more than 3 times those in the medical field" (Duhart, 2001).*

Violent assaults most of the times happen in specific circumstances: "during admission, change of shift, mealtimes, visiting hours, and any period of change when nursing staff have varied tasks to complete" (Barlow, Grenyer and Ilkiw-Lavalle, 2001). Dack et al. (2013) have recently conducted a review and meta-analysis showing wide oscillations of the figures that describe aggressive behaviours of patients in inpatient settings, varying from 8% to 44% according to different sources in the literature. According to the authors, in inpatient wards around 41% of the clinical staff has experienced violent or threatening behaviours, the rate for nursing staff is nearly 80%. In particular, the impact on mental health nurses can be significantly higher, exposing them to higher risk of injury and affecting their morale. While for inpatients, the authors reported that approximately a third of them has experienced violent behaviours during their psychiatric care (Dack et al, 2013).

Several reasons have been proposed in the literature to explain the variation in reported rates of violence: (1) Inconsistency in the definition of violence and aggression (Cutcliffe & Riahi, 2013a), (2) Different methods for collecting data, (3) Actual variation of violence in psychiatric wards, and (4) Underreporting of incidents (Iozzino et al., 2015). However, despite the variation in estimates, there is a unanimous agreement on the prevalence of violence in the mental health sector, with a significant number of health care professionals facing at least one assault in their career (McDermott & Holoyda, 2014). Thus, many studies have focused on how to guarantee the safety of both staff and patients. Understanding the drivers of aggression and violence in a psychiatric unit are one starting point to better understand this complex phenomenon.

Inpatient aggression and violence can significantly decrease the staff morale, affecting the therapeutic progresses as well as the utilization of hospital resources (Serper et al., 2005; Neufeld, Perlman and Hirdes, 2012). It can lead to increased absenteeism and turnover (Foster et al., 2007; Nijman et al., 2005a; 2005b), reduced levels of motivation (Arnetz & Arnetz, 2001; Hahn et al. 2010; Needham et al., 2005), fear based behaviours towards patients who present aggressive/violent behaviours (Duxbury, 2002), and a tendency to reduce the interaction with patients (Luckhoff et al., 2013; Michie & West, 2004). Consequently, the perceived satisfaction of patients is impacted resulting in lower satisfaction scores (Abderhalden et al. 2006; Meehan et al., 2000), with associated higher costs for running the facility (LeBel & Goldstein, 2005).

2. Definition of violence and aggression

There is inconsistency in the literature as to the definition of violence and aggression as well as what constitutes a violent incident (Cutcliffe & Riahi, 2013; Bjørkly 2006). The terms have also been used interchangeably, although they relate to different conditions (Littrell & Littrell, 1998). It is thus important to agree on common definitions for these terms.

2.1 Violence

Violence can be defined as an explicit eruption of physical force which "abuses, injures, or harms another individual or object" (Littrell & Littrell, 1998).

Violence can also refer to a mental attitude or can describe an interpersonal relationship (Rocca, Villari & Bogetto, 2006). The WHO has classified violence into three categories, depending on the perpetrators:

1. Self-directed violence
2. Interpersonal violence
3. Collective violence

2.2 Aggression

The term aggression has a broader breadth. Bandura (1973) compared doing research in this field to navigating into a sort of semantic labyrinth of concepts embracing a wide spectrum of phenomena and activities (Bandura, 1973). What makes studying human aggression challenging is the fact that it occurs occasionally, and it is often not acknowledged nor reported (Bjørkly, 2006). Aggression has been linked to anger. Panksepp (1998) has argued that anger is related to the efforts of the individual in pursuing ongoing desires and success in the competition for resources. When anger is chronic, it can be severely debilitating for the individual. Aggression can be a verbal or physical attack and can be "either appropriate (e.g., self-protective) or, alternatively, it may be destructive to the self and others" (Liu, 2004).

There are different definitions of injury in the literature as well as among institutions. The Merriam-Webster defines aggression as

> "*a forceful action or procedure (such as an unprovoked attack) especially when intended to dominate or master; the practice of making attacks or encroachments; hostile, injurious, or destructive behavior or outlook especially when caused by frustration*" (Merriam-Webster Dictionary, 2018).

According to Whitley et al. (1996), an important point of reference is the standard definition used in reporting employee injuries to the Occupational Safety and Health administration (OSHA) which comprises "lost work days, loss of consciousness, restriction of work or motion, termination of employment, transfer to another job, or medical treatment other than first aid" (Whitley et al., 1996). As noted by the authors, the OSHA standards do not take into consideration many episodes of violence and assaults suffered by nursing staff.

Aggression is a construct that presents a high degree of complexity, and can be categorized in different ways (Siever, 2008), such as considering the mode of aggression (e.g., physical/verbal, direct/indirect), who is targeted (e.g., others/self-directed), or the cause of aggression (e.g., medical factors, psychiatric disorders, etc). The most widely adopted classification of aggression has been the impulsive/instrumental dichotomy (Rosell & Siever, 2017). The "instrumental" type has also been named proactive, premeditated or predatory, whereas the "impulsive" type can be also defined as reactive, affective or hostile (Siever, 2008). The impulsive type refers to the immediate reaction of the individual to a perceived threat or provocation, with high levels of autonomic arousal; the instrumental type is characterized by a premeditated, goal-oriented plan, not necessarily showing signs of autonomic arousal (Siever, 2008; Rosell & Siever, 2017). Impulsive aggression is also referred to as "hot-headed" response, while "cold-blooded" response is used for the instrumental one (Bushman & Anderson, 2001). According to this dichotomous view, there are at least three elements that define the difference between impulsive and instrumental aggression:

I. **Primary goal of aggressive act**: The primary goal of impulsive aggression is to harm someone, whereas for instrumental aggression harm is a "means to some other end" (Bushman & Anderson, 2001). Since aggression can be motivated by a several different goals, a significant manifestation of this variant can be interpreted as instrumental aggression (Tedeschi & Felson, 1994).

II. **Presence of anger**: Anger is always present in impulsive aggression, in contrast to instrumental aggression. However, this dichotomy might pose some difficulties in categorizing acts of violence that result from acts of revenge, where the initial cold determination can be rooted in anger (Bushman & Anderson, 2001).

III. **Capacity and complexity of planning and calculation**: Impulsive aggression is characterized by unplanned, impulsive behaviour, where the individuals perpetrating the act barely consider the consequences of their actions. Conversely, in the premeditated aggression the individuals have a mental plan with different alternatives, and typically calculate, at least at a basic level, costs and benefits. Not all the behaviours can fit into the hot/cold categories: some of them are "warm", and we probably need to clarify how much premeditation and planning are needed to consider an aggressive act as instrumental (Bushman & Anderson, 2001).

An alternative to the impulsive/instrumental dichotomy is the reactive/proactive classification, validated initially in children and adolescents, and more recently in adults (Rosell & Siever, 2017). Proactive aggression can be associated with the instrumental aggression, and identifies behaviours controlled by external rewards (Cima et al., 2013). Reactive aggression can be associated with the impulsive type, and typically refers to unplanned, hostile reactions to frustration or to stimuli perceived as a threat (Cima et al., 2013). This framework assumes that the two subtypes, reactive and proactive, coexist, are highly intercorrelated and can be both assessed dimensionally (Rosell & Siever, 2017). Both variants have been assessed in adults diagnosed with borderline personality disorder (Gardner, Archer & Jackson, 2012), as well as antisocial personality disorder (Lobbestael, Cima & Arntz, 2013).

Violence and aggression are multifactorial constructs, depending on the complex interaction of biologic, psychologic and social variables. Currently, two major systems are considered to be responsible for many aspects of aggressive and violent behaviour: the limbic system and the frontal lobes (Rocca, Villari & Bogetto, 2006). The functional and neurobiological findings related to aggression and violence, as well as the main psychological theories adopted to explain them, are discussed in Chapter 3.

3. Complexity of violence in inpatient settings

The phenomenon of violence in inpatient wards is complex, and attributing the responsibility for it exclusively to the patients would be misleading. The number of studies in this field has grown considerably in the last two decades. There is evidence that clinicians tend to hold the patient responsible for the problem, and rarely reflect about the impact of their actions/behaviours (Cutcliffe & Riahi, 2013 a; Duxbury & Whittington 2005). In this regard, an interesting research question is to study mental health issues from the perspectives of clinical staff as well as patients. The latter perspective is barely taken into account in the literature (Whittington, 2000). This is especially true if we consider the point of view of aggressive patients, since the majority of interview studies have been centred on mental health professional's views (Duxbury & Whittington 2005; Marangos-Frost & Wells 2000).

Quintal (2002) has noticed that in the literature violent behaviours do not usually appear to be isolated acts. Rather, we typically assist to an escalation of agitation and violence, from a calm state to a state of violence. Littrell and Littrell (1998) have defined a linear progression to show the various gradations of violence. This progression shows five states: calm, anxious, agitated, aggressive and violent. Agitation, described as an offensive verbal, vocal or motoric activity, is considered the less threatening state; however, if not treated properly, it might lead to various forms of aggression and ultimately to explicit acts of violence (Littrell and Littrell, 1998).

4. External and internal factors

Due to its complexity, the phenomenon of inpatient violence has been analysed from different points of view. In fact, a violent incident often results from the combination of different factors, external and internal. The internal factors are related to the individual characteristics and the state at the moment of the incident of both the acting subject and the victim. These elements are combined with a set of external factors, represented by the characteristics of the environment and the set of circumstances (Davison, 2005). Consequently, when we consider the risk factors associated with violence, we can group them in three main categories: internal patient factors, internal victim factors and external environmental factors.

4.1 Environmental predictors of violence

Many service users believe that external factors play a stronger role as precipitants of inpatient violence than their counterpart (National Institute for Clinical Excellence, 2004). Thus, manipulating the environment is one strategy to decrease the risk of violence.

Recently, several studies have pointed out that there are characteristics that might significantly impact the manifestation of violence. Among the most important environmental factors associated with an increase in the risk of violence is a high use of temporary staff, lack of privacy and overcrowded environments, poor levels of interaction between patients and the staff, and a lack of structured activities (Royal College of Psychiatrists, 1998). A study named "City-128" had the purpose to assess the impact of different factors (such as physical containment, staff demography and behaviours, etc.) to rates of self-harm and suicide within psychiatric units (Bowers, 2007). The usage of coercive measures and locked doors were associated with more self-harm, whereas intermittent observation was associated with a decrease in self-harm episodes. It was concluded that, in order to reduce self-harm in inpatient wards, it would be recommended to shift towards "a more liberal psychiatry with more emphasis on patient freedom and responsibility, and less use of containment" (Bowers, 2007). According to a qualitative study conducted by Haglund et al. (2006), the disadvantages of locked doors seem to be significantly higher than the advantages. The authors found eight categories of advantages and eighteen categories of disadvantages. The advantages mentioned by nurses and mental health nurse assistants refer to having a more efficient control over patients and delivering a better care. The most important disadvantages mentioned by nurses and mental health nurse assistants referred to a feeling of confinement in patients, extra work for the staff, more emotional problems for patients accompanied by a feeling of being dependent, making more explicit the power of staff and forcing some patients to adapt to the needs of other patients (Haglund et al., 2006).

Finally, it is important to underline the relationship that exists between living in a socially disadvantaged environment and deviant behaviours. The social disorganization theory is one of the most used framework to analyse the relationship between criminal behaviours and living in a disadvantaged neighbourhood (Silver, 2000). This model is particularly considerably more relevant today. In fact, nowadays the number of psychiatric patients admitted for long term periods is significantly smaller than it used to be prior to the 1960s, thanks to the implementation of several policies of deinstitutionalization (Silver, 2000). This theory suggests that communities having high levels of social disorganization do not provide an adequate social support to residents. As a result, if a psychiatric patient is discharged in a socially disadvantaged neighbourhood, the probability to engage in violence against other individuals would be significantly higher than living in a more stable and less disorganized environment (Silver, 2000).

4.2. Individual predictors of violence

Among the internal factors influencing violence in inpatient settings are demographic and personal characteristics. There is sufficient evidence that shows that people experiencing pain have a limited capacity to tolerate frustration and stress, and are more prone to react with aggressive or violent behaviours (Cutcliffe & Riahi, 2013a). Serper et al. (2005) demonstrated that aggressive behaviours of patients are affected by executive dysfunction. According to the authors, patients with executive dysfunction and psychosis suffer from a deficit in behavioural regulation and control of impulses. Consequently, patients with acute symptomatology and executive dysfunction are considered at high risk of committing aggressive behaviours. The results from this study stress the importance of using specific neuropsychological testing to predict violence in psychiatric units (Serper et al., 2005). In addition, many studies have reported a significant association between both traumatic experiences and substance addiction and violent/aggressive behaviours (Cutcliffe & Riahi, 2013a). Not surprisingly, different substance abuse treatments are now considered a key component to mitigate violent behaviours in individuals with mental illness (Hiday, 2006). Psychosis is also an element particularly relevant in this context, since its association with violence has been long debated. A significant association between psychosis and violence has emerged in a meta-analysis conducted by Douglas et al. (2009) who considered 204 studies. In their study, the role of psychosis as a risk factor is intermediate if compared to other factors. Larger risk predictors are the antisocial personality disorder, or early-onset criminal behaviour (Douglas et al., 2009).

Johnson (2004) categorized the factors associated with aggression in four groups: staff-related variables, unit-related variables, interactional variables, and patient-related variables. The author concludes that assessing specific symptoms or behaviour clusters as well as considering clinical variables are more reliable strategies to predict violence, compared to dispositional factors (Johnson, 2004; Woods & Ashley, 2007). Other factors acting as predictors of violence include an early onset of psychiatric illness, history of alcohol abuse, past characterized by victimization, and poor premorbid adjustment. A poor rapport with the staff and the admission status seem also to play a role (Johnson, 2004). Some authors demonstrated an association between violent behaviour and involuntary admission (Johnson, 2004; Serper et al., 2005) as well as admission to a locked unit (Johnson, 2004). Different authors have also reported compulsory admission among the factors associated with aggressive behaviour (Woods and Ashley, 2007; Soliman & Reza, 2001). Patients suffering from a more severe psychopathology are more prone to be aggressive (Arango et al., 1999; Johnson, 2004). The most consistent results link a diagnosis of psychosis, schizophrenia, mania, and a spectrum of organic syndromes. However, the relationship between both age and gender with inpatient violence showed inconsistent results (Johnson, 2004). Similarly, Otto (2000) and Anderson (2004) have concluded that gender should not be considered as a risk factor for inpatient violence. A different result was reported by Serper et al. (2005), who found that women are more aggressive during hospitalization. Some have suggested that the higher rate of aggression in hospitalized women might be due to a bias where women who commit aggression are considered threatening since otherwise women are usually perceived to be less dangerous by the staff.

Anderson et al. (2004) have organized the patient-related risk factors in four categories: dispositional, contextual, clinical and historical. Dispositional factors are mainly descriptive and static risk factors such as demographic, personality, cognitive. In this category age appears to be a risk factor for both individuals with mental illness and individuals without, showing that the category more at risk is represented by persons in their late teens and early twenties (Anderson et al., 2004). The historical factors are also very relevant, especially when young patients are exposed to violence and engage in experiences that reinforce their display of violence (Jenkins & Bell, 1997; Anderson et al., 2004). Among the contextual risk factors, we can find stress, social support, substance use, or being victim of assault. Substance abuse and stress are associated with an increased incidence of violence. Clinical factors refer to the diagnosis and the assessment of symptoms and functioning. Receiving a diagnosis of substance abuse or dependence has shown a significant association with increased risk of violence (Anderson et al., 2004; Swanson, Holzer, Ganzu et al., 1990). The association of substance abuse and inpatient violence has been confirmed also by other studies (Serper et al., 2005; Soliman & Reza, 2001).

Several personality disorders have been associated with violent behaviour in various contexts, the degree of psychopathy in these patients could also impact the potential for violence (Anderson et al., 2000). In addition, certain symptoms typical of the schizophrenic disorders might expose patients to engage in violent behaviours, rather than the diagnosis of schizophrenia per se (Woods and Ashley, 2007; Anderson, 2004).

Steinert and colleagues (2000) found a significant association between the psychopathological status at admission and aggressive behaviours during the permanence in a psychiatric facility. The authors concluded that aggressive behaviour in inpatient settings was associated "more with general aspects of disturbance such as thought disorders and hostility", rather than with specific psychotic features (Steinert et al., 2000). Chou et al. (2002) have studied assaultive behaviour and assault in acute inpatient wards, by examining a total of 855 episodes of violence from 287 patients. They concluded that the factors associated with assaultive episodes and behaviour are previous history of aggressive behaviour, a diagnosis of psychotic disorder, duration of admission and history of smoking. In addition, there are some contributing factors patient/nurse ratio and space density, and staff factors such as age, training and length of work experience (Chou et al., 2002).

We summarise our findings in Table 1 and 2. Table 1 lists the different predictors that have been studies in the literature. In Table 2 we classify these predictors according to whether they are individual or environmental.

PREDICTORS	**AUTHORS/STUDIES**
<u>Precipitants of inpatient violence</u> • High use of temporary staff • Lack of privacy • Overcrowded environments • Poor level of interaction between patients and staff • Lack of structured activities	Royal College of Psychiatrists (1998)

<u>Precipitants of self-harm</u> Coercive measure and locked doors	Bowers (2007)
<u>Precipitants of inpatient aggression and violence</u> • Limited capacity to tolerate frustration and stress • Experiencing pain • Traumatic experiences • Substance addiction	Cutcliffe and Riahi (2013a)
<u>Precipitants of inpatient aggression and violence</u> Executive dysfunction and psychosis	Serper at al. (2005)
<u>Factors associated to inpatient violence and aggression</u> • Psychosis (intermediate) • Antisocial personality disorder (higher) • Early-onset criminal behaviour (higher)	Douglas et al. (2009)
<u>Factors associated to inpatient violence and aggression</u> • Psychiatric illness • History of alcohol abuse • Past characterized by victimization • Poor premorbid adjustment	Johnson (2004) Woods, P. and Ashley, C. (2007)
<u>Factors associated to inpatient violence and aggression</u> • Poor rapport with the staff • Involuntary admission	Johnson M.E. (2004) Serper at al. (2005)
<u>Patient-related risk factors, four categories:</u> 1. Dispositional (demographic, personality, cognitive etc.) 2. Contextual (stress, social support, substance use, being victim of assault) 3. Clinical	Anderson et al. (2004)

4. Historical (being exposed to violence and engage in experiences that reinforce violence)	
Factors associated to inpatient violence and aggression • Several personality disorders • Specific symptoms typical of the schizophrenic disorders	Woods and Ashley (2007) Anderson et al. (2004)
Factors associated to inpatient violence and aggression • history of aggressive behaviour • diagnosis of psychotic disorder • duration of admission • history of smoking	Chou et al. (2002)

Table 1. Predictors of aggression in the literature.

INDIVIDUAL PREDICTORS	**ENVIRONMENTAL PREDICTORS**
• Limited capacity to tolerate frustration and stress • Experiencing pain • Traumatic experiences • Substance addiction	• High use of temporary staff • Lack of privacy • Overcrowded environments • Poor level of interaction between patients and staff • Lack of structured activities
• Executive dysfunction and psychosis	• Coercive measures • Locked doors
• Psychosis (intermediate) • Antisocial personality disorder (higher) • Early-onset criminal behaviour (higher)	• Poor rapport with the staff • Involuntary admission • Duration of admission

• Psychiatric illness • History of alcohol abuse • Past characterized by victimization • Poor premorbid adjustment • Several personality disorders • Specific symptoms typical of the schizophrenic disorders • History of aggressive behaviour • Diagnosis of psychotic disorder	

Table 2. Individual and Environmental Predictors of Aggression

5. Management of violence in inpatient settings

To manage incidents of aggression and violence, health care professionals need to use strategies calibrated to the risk posed to self and others (DoH, 2005). The methods used to manage violence can be generally classified as non-coercive or coercive. Non-coercive methods are more collaborative approaches, whose aim is to calm down patients using different techniques, giving them the possibility to elaborate anger and frustration excluding any form of violence (Davison, 2005). Examples of these methods are:

- De-escalation, described by The National Institute for Health and Care Excellence (NICE) as “talking with an angry or agitated service user in such a way that violence is averted and the person regains a sense of calm and self-control" (NICE 2015, p 30). De-escalation methods include both verbal techniques such as using a calm tone of voice and avoiding to threaten the patients, and non-verbal techniques such as awareness of self, eye contact or body stance (Spencer & Johnson, 2016; Cowin, 2003; Johnson 2011). De-escalation is recommended by the NICE as an early intervention that can prevent harm caused by escalation to the crisis phase (NICE, 2015). De-escalation techniques are helpful to redirect patients towards a "calmer personal space" (Cowin, 2003).

- Time out, which consists in obtaining the consensus from a patient to stay in a room, usually their bedroom, until they calm down (Bowers et al., 2012). Time out can substitute the use of geographical restraint, as a less coercive practice.
- Observation, an intervention in which a member of the staff "maintains contact with a service user to ensure the service user's safety and the safety of others" (NICE, 2015).

Coercive methods are necessary when the situation is acutely dangerous or in case of failure of more collaborative approaches (Davison, 2005). These methods are:

- Physical restraint, practiced by trained staff who use recognised techniques to hold a patient and limit their capacity to move (Davison, 2005). Proper training is extremely important to prevent injuries to all the individuals involved. The NICE guidelines admit the use of pain in some physical restraint techniques, but they should be adopted only under exceptional circumstances (Davison, 2005).

- Geographical restraint (or seclusion), a method used for physical containment consisting in moving a patient to a quieter place. A common practice in Psychiatric Intensive Care Units (PICUs) is the utilization of seclusion suites, where patients are moved to manage a variety of behaviours, such as disruptive and aggressive behaviour, acute psychiatric symptoms or self harm (Spencer & Johnson, 2016). The use of seclusion is governed by strict guidelines, as it involves restriction of a patient's liberty, and it is recommended only as a last resort (Davison, 2005).
- Chemical restraint, also called rapid tranquillisation, reduces the arousal of a patient using medications, such as intramuscular injections. It aims to prevent violence and limit the immediate suffering of the individual (Davison, 2005).

6. Conclusion

Research and practice have demonstrated that reducing incidents due to violence and aggression in inpatient setting is possible. However, it requires that facilities have enough resources dedicated to monitoring, management and prevention strategies, and the staff receives adequate training covering all the techniques and methods mentioned above. Non-coercive and more collaborative methods should be preferred, whenever possible.

There is sufficient evidence in the literature which demonstrates how clinicians, through their attitude and beliefs, influence the issues related to aggression and violence in mental health care. Receiving a specific training for violence has demonstrated to be a key element in improving the self-confidence of the clinical staff. Bjorkdahl et al. (2012) have studied the impact of training staff on violence prevention and management. One such training program is called the Bergen model that is developed from City theoretical model (Bowers, 2002). The City model identifies three factors that appear to be critical in reducing conflicts and violence in psychiatric wards: “positive appreciation of patients, emotional regulation and effective structure” (Bjorkdahl et al., 2012). According to the Bergen model, the front line of prevention is represented by primary prevention factors that promote a positive relationship between staff and patients. The secondary and tertiary sections of the training address different aspects such as “limit-setting styles and negotiation, self-defence, physical restraint techniques (so-called pain compliance techniques are not used) and safety issues, the use of mechanical restraint, seclusion and forced medication, post-incident sessions with the patient and with the staff and critical reviewing of violent incidents” (Bjorkdahl et al., 2012). The authors conclude that the training model had a positive influence on preventing the violence and promoting a better management climate within a psychiatric ward.

On the other hand, it is important to analyse the relationship between violence and the mental health-care as a system (Cutcliffe & Riahi, 2013b). Factors such as culture, policies and practices in place play a key role. It is possible to analyse these elements considering different levels: we might analyse policies and practices from a governmental point of view; we might concentrate on a specific institution or hospital; we can narrow our focus and study a single ward in a psychiatric facility (Cutcliffe & Riahi, 2013b). One of the most important goals of an inpatient unit is the implementation of procedures and protocols to guarantee both patients' and staff's safety, the delivery of comprehensive assessments of patients' problems, the delivery of physical care and treatments, and the capacity to understand the basic self-care needs of patients. An acute ward can be defined efficient if patients are safe, properly assessed and treated. In addition, an efficient acute ward can accomplish the goals established meeting reasonable deadlines and keeping the costs minimal (Dack et al., 2013; Bowers et al., 2005).

Chapter III

Psychology and neurobiology of aggression

Aggression and violence are complex constructs with multiple determinants - not only biological, psychological, and social factors, but also political, socioeconomic and cultural factors (Siever, 2008; Rosell & Siever, 2017). They are studied in several disciplines, each one of them adopting a different methodological framework to address the complexity that these phenomena represent (Bjørkly, 2006).

The present chapter will provide an overview of the main psychological theories developed to explain violence and aggression, followed by an overview of the main structural and functional neuroanatomical findings related to aggression and violence in human beings.

Psychological Theories of Aggression

This chapter will provide an overview of four psychological theories of aggression. In line with Bjørkly (Bjørkly, 2006), the first three theories of aggression refer to the psychoanalytic theory, the drive theory and the social learning theory. The fourth theory is the General Aggression Model (Anderson & Bushman, 2002), developed recently to integrate different domain-specific theories.

1. The Psychoanalytical Theory

The psychoanalytic theory postulates that human behaviour originates from two basic and opposite instincts: Eros, the life instinct, and Thanatos, the death instinct. This dual-instinct theory is characterized by a polarized relationship between these two forces (Freud, 1920). According to Freud, feelings of anger and hostility result from conflict and unconscious guilt, following the same dynamic that is in place for sexual wishes. The nature of the relationship between the two forces mentioned, either a conflicting or a harmonic relationship, is the foundation for explaining the how and why several impulses containing both sexual and aggressive components originate, including a variety of clinical manifestations (Bjørkly, 2006). The key point about the role of the death instinct refers to the need to direct the aggressive force to the external environment, and thus avoiding self-destruction. In fact, according to Freud, the equilibrium within the individual can be reached when the aggressive impulses are mitigated or fused with love. There are some events that can interfere with this process of fusion between love and aggression, such as traumas or abuses, causing the increase of the destructive energy which leads to destructive behaviours. Freud postulated the existence of a mechanism through which the individual can reduce the tension created by the accumulation of the destructive energy: the catharsis. According to the psychoanalytic theory, the death instinct (or Thanatos) is the basis of aggression. As noted by Bjørkly (2006), the definition of aggression within the psychoanalytic framework is wide, embracing a significant variety of manifestations. Without the role played by the death instinct, the aggressive forces of the individual would be directed internally causing the self-destruction. The death instinct “pushes” the aggressiveness of the individual towards the external environment. One of the main objectives in psychoanalysis is to help patients in understanding the unconscious motivations at the base of their behaviour, including the comprehension of what psychological mechanisms are involved in aggressive manifestations.

Pioneered by Freud, the psychoanalytical model has faced considerable changes over the course of the last century, with implications for clinical practice. Consequently, our understanding of aggression has changed from being more focused on the single individual and the intrapsychological processes involved in the aggression, to a more interactional view where the phenomenon of the aggression can be understood by analysing the interaction of the individual with the social environment. The attachment theory, formulated by John Bowlby, represents the opposite pole of the spectrum. The attachment theory preserves some of the key concepts theorized by Freud in regards to the importance of the early experience, the defensive processes, and close relationships. Bowlby’s model focuses on infants and children who are viewed to be more competent and environmentally oriented. In addition, the relationship between the children and their caregivers plays a key role in the learning process about the environment (Waters et al., 2005). The concept of relationship is one of the pillars of the attachment theory. Individuals are considered social creatures, and, thus, the capacity to develop social relationships is crucial for the survival of the individual. In addition, among all the social relationships, a special role is played by the relationships formed between the child and his or her parental figures. The theoretical framework created by Bowlby stresses the existence of a behavioural control system constructed over time, resulting from the behavioural interactions between the infant and the caregivers. The experiences made in the early stages of life are supposed to influence and shape the future relationships (Waters et al., 2005). Modern psychoanalytic theories stress the importance of cognitive functions and social interactions in the comprehension of human aggression (Bjørkly, 2006), theorizing the construction of internal cognitive models based on the interaction between the individual and the attachment figures.

2. The Frustration-Aggression Hypothesis

The theory of aggression formulated by Freud has been harshly criticized. The lack of an empirical basis to test and prove the theory was one of the main critiques. Thus, in 1939 a group of researchers formulated a new theory of aggression and named it "the frustration-aggression hypothesis" (Dollard, Doob, Miller, Mowrer, and Sears, 1939). The authors defined frustration as "that condition which exists when a goal response suffers interference" (Dollard et al., 1939). When goal-oriented behaviours find obstacles in their pathway, it gives rise to frustration. The authors assumed that when individuals cannot accomplish their goals they become frustrated, and respond to this state with aggressive behaviours. As a consequence, aggression is theorized as the reaction to the frustration state, a consequence of a reaction to an external stimulus. They postulated that frustration always leads to some form of aggression, and a frustrating stimulus is necessary to start developing an aggressive response (Dollard et al., 1939). In the development of an aggressive response, both the genetic and the environmental component are thought to be equally determinant (Bjorkly, 2006). Thus, one of the major differences between the drive theory and the Freudian theory of aggression refers to the nature of the aggressive "drive", which is external and not internal as theorized by Freud (Dollard et al., 1939). However, individuals can learn to mitigate the aggressive forces when they start experiencing a state of frustration. In clinical practice, the theoretical framework described by the frustration-aggression hypothesis can be helpful to analyse some of the factors contributing to the escalation of violence in psychiatric wards. For example, different studies have shown a clear link that involves communication problems between patients and clinical staff and the risk of violence and aggression. Whittington and Wykes (1996) studied the strategies adopted by the staff in psychiatric setting to cope with different levels of anxiety due to aggressions committed by patients. Two strategies are particularly interesting: escape/avoidance and confrontive coping. These approaches are essentially based on increasing the amount of interaction with patients (confrontive coping) or decreasing it (escape/avoidance). The confrontive coping results in a "deterioration in the quality of such interaction" (Whittington & Wykes, 1996), with the staff displaying hostile feelings and behaviours. Conversely, the strategy of avoiding interaction would prevent staff behaviours that might elicit violent reactions from psychiatric patients. Staff who have already experienced assaults might feel powerless and have a higher risk of engaging in risky behaviours with the aim to increase their sense of control, resulting in a vicious circle (Whittington & Wykes, 1996). The frustration, caused by the adoption of a confronting coping, will in turn lead to aggressive episodes. Other elements that have a relationship with frustration in psychiatric patients are related to the architecture of psychiatric wards. Ng et al. (2001) have found that the variables “density”, “privacy” and “control” are strictly interrelated and can help in understanding the impact of architecture on human mood and behaviour. They have found that “a sense of frustration and anger are expected outcomes, especially when they [patients] feel that they have lost control over their environment” (Ng et al., 2001). Consistent with the frustration-aggression model, a high population within a ward is the external stimuli that can cause frustration in patients, who will negatively react with violent behaviours. Björkdahl et al. (2010) studied the approaches adopted by nurses towards patients in acute psychiatric care. They distinguished between two approaches: the bulldozer, where wards are kept ordered, and the ballet dancer, where the focus is on building a relationship with patients. There is a risk of engaging in harmful and uncaring actions when nurses adopt the bulldozer approach (Björkdahl, Palmstierna, & Hansebo, 2010).

3. Social Learning Theory

The Social learning theory (SLT) can be considered a general theory of human behaviour. It tries to explain how the individual develops the specific competencies, often identified as “personal”, within the social context where the process of learning takes place. The most important and influential author who contributed to the development of the social learning theory is certainly Albert Bandura. Bandura’s work has been instrumental for the SLT theory in studying aggression in human behaviour. Using

Bandura's words "human functioning...involves interrelated control systems in which behaviour is determined by external stimulus events, by internal processing systems and regulatory codes, and by reinforcing response response-feedback processes" (Bandura, 1969, p. 19). Bandura released the first complete and coherent theory focused on modifying deviant behaviour (Bandura, 1969). The behaviour is the result of several learning processes that take place in a setting of mutual and reciprocal interactions between the individual and the environment, where the individual is an active organism that both influences and is influenced by the environment. Bandura (1969) defined reciprocal determinisms where he argued that behaviour, environment and cognitive processes are strongly interrelated and influence each other. The SLT model is clearly detached from the two views that dominated the American psychology during more than two decades, the psychoanalytic and associative conditioning-learning models. The application of the social learning principles to the investigation of aggressive behaviour in humans has been pioneered initially by Arnold Buss and Albert Bandura (Bjørkly, 2006). Bandura (19979) defined aggression as a

> *"behaviour that results in personal injury and physical destruction. The injury may be physical, or it may involve psychological impairment through disparagement and abusive exercise of coercive power...Aggression refers to complex events that include not only injurious behaviour but judgmental factors that lead people to attach aggression labels to some forms of harmful conduct but not to others...The same harmful act is perceived differently depending on sex,*

age, attractiveness, status, socioeconomic level, and ethnic background of the performer".

The author argued that harmful acts of favoured individual and groups are perceived differently (Bandura, 1979). In Bandura's view, an exhaustive theory of aggression must describe in detail three major components within the SLT framework. The first component is called the "origins of aggressions". It defines how aggressive behavioural patterns originates and how they are developed. It constitutes three elements: observational learning, reinforced performance and structural determinants. The second component is called the "instigators of aggression", and defines what factors act as triggers of aggressive behaviours. Modeling influences (e.g., disinhibitory or facilitative), aversive treatment (e.g., physical assaults or adverse reductions in reinforcement), and instructional control are some of the elements that act as instigators. Finally, the third component is defined as the "regulators of aggression". It focuses on the mechanisms that sustain aggressive actions after they have been initiated. It involves the external reinforcement (such as social and status rewards), the punishment (inhibitory or informative), the vicarious reinforcement and the self-reinforcement (self-reward or self-punishment).

SLT has been criticized as not being a specific theory of aggression. Rather, it proposes a structure of learning principles that regulate different behaviours, including the aggressive ones. Some authors have criticized the fact that this theory is strongly limited to laboratory settings and thus underestimating the impact that social interactions have on human behaviour when individuals interact with each other. Others have criticized the lack of clarity of STL in explaining the mechanisms involved in modeling effects in aggressive behaviour (Bjørkly, 2006).

SLT principles find different applications in modern clinical psychology. In fact, there are many points of contact between SLT and the cognitive behaviour therapy (CTB). These two "approaches" share the foundations of behavioural principles. These include the processes that govern how human beings learn, as well as the critical roles played by both external consequences and self-regulatory mechanisms in regulating the human behaviour (Bjørkly, 2006). The principles that constitute the foundation of SLT find their counterpart in clinical settings, confirming the clinical relevance of Bandura's theory.

4. The General Aggression Model

The General Aggression Model (GAM; Anderson and Bushman, 2002) is the most recent theory of aggression. It is a comprehensive, integrative framework that embraces different domains, to generate a biosocial-cognitive model capable of addressing short-term and long-term effects of several variables that play a role in human aggression (Warburton & Anderson, 2015). This model combines five recent theories of aggression (Anderson and Bushman, 2002), which overlap considerably:

1. the Cognitive Neoassociation Theory (Berkowitz 1989, 1990, 1993)

2. the Social Learning Theory (Bandura, 1979, 2001)
3. the Script Theory (Huesmann 1986, 1998)
4. the Excitation Transfer Theory (Zillmann 1983)
5. the Social Interaction Theory (Tedeschi & Felson 1994)

The Cognitive Neoassociation Theory (Berkowitz 1989, 1990, 1993) is grounded in the emerging field of neural connectivity. It postulates that aversive events (such as frustrations and provocations) produce a negative affect, which in turn evokes emotions, memories, thoughts and behaviours associated with the fight-or-flight reactions (Allen & Anderson, 2017; Warburton & Anderson, 2015; Anderson & Bushman, 2002). The fight associations are linked to reactions based on feelings of hostility, aggressivity and anger; flight associations evoke feelings of fear, pushing the individual to avoid the situation. All the associations are strictly interconnected, forming a network of memory structures (Allen & Anderson, 2017). Moreover, the theory postulates the existence of a neural process, that elicit the association of cues present during an aversive event with the event itself and with the cognitive and emotional responses elicited (Anderson & Bushman, 2002). This model can be viewed as reformulated version of the frustration-aggression hypothesis, and it is helpful to address situations characterized, for example, by reactive aggression (Allen & Anderson, 2017).

The Script Theory (Huesmann 1986, 1998) introduces the element of the "scripts", concepts that are highly associated in memory that describe a specific situation, providing a sort of behavioural guide related to that situation. Scripts can be acquired through direct learning or observation. According to this theory, scripts are strongly interconnected. The more a person has access to certain scripts, such as those associated with aggressive behaviours, the easier the access to these scripts will become. As a result, these scripts can be become automatic and even generalized to other circumstances, increasing the probability of behaving according to those scripts (Warburton & Anderson, 2015). Given its nature, the Script Theory provides more details on the comprehension of how social learning processes function in a variety of situations, and how automatic behaviours are formed and acted.

The Excitation Transfer Theory (Zillmann 1983) lies in the field of psychophysiology, and considers the concept of physiological arousal as a key determinant in understanding aggression. The starting point is the slow and gradual dissipation of the physiological arousal. When an arousing event occurs and the window of time that separates it from a second arousing event is not long enough, the residual arousal from the first event can be misattributed to the second event. In cases when the second event evokes hostility or anger, the second arousal will exacerbate the feelings of hostility and anger (Allen & Anderson, 2017; Warburton & Anderson, 2015; Anderson & Bushman, 2002). In addition, individuals who label themselves as angry can be hostile for long periods of time, going far beyond the dissipation of the initial arousal.

The Social Interaction Theory (Tedeschi & Felson 1994) considers aggression as a form of social influence, driven by expectations, rewards, costs and probability to succeed. Given its nature, this theory is particularly useful in explaining aggressive behaviours motivated by the necessity to protect the individual's self-esteem, or to accomplish higher goals (Allen & Anderson, 2017).

The effort to combine the aforementioned theories into one model resulted in a framework that aims at explaining the complex interaction of social, cognitive, developmental and biological factors. The first aspect that differentiates GAM relates to the importance of the knowledge structures, which every individual develops through experience. Knowledge structures influence the human aggression in many aspects since they impact not only the perception processes, but also interpretation processes, decision making, and behaviours (Allen, Anderson & Bushman, 2018).

GAM's focus is on the "person in situation" (Anderson and Bushman, 2002), or an episode, which consists of a cycle of interactions in a social environment, defined by specific processes. GAM can be separated in two main parts: proximate and distal processes (see Figure 1). The proximate processes are used to explain aggression by using three stages: inputs, routes, and outcomes. The first stage explains the interaction of a person and situation factors, and how this interaction can ultimately increase or decrease human aggression. Inputs refer to the factors that constitute every act of aggression, and can be divided into person factors and situation factors. The individual inputs, or person factors, address the individual differences that play a role in how the person responds to a situation. They include biological factors as well as attitudes. The environmental factors, on the other hand, relate to elements or characteristics of the situation that might influence the aggressive response, and can include cognitive cues as well as provocation stimuli.

The second stage addresses the three main modalities, or routes, through which person and situation factors act to influence appraisal and decision processes: affect, cognition, and arousal (Allen, Anderson & Bushman, 2018). Changes in these three domains reflect changes in a person's internal state, which in turn will affect the likelihood of violence and aggression. The model illustrated in Figure 1 explains the reciprocal interactions between internal variables and between the three domains. The variables involved in these routes in turn influence the appraisal and decision-making processes.

The third stage addresses how an individual evaluates and reacts to a certain situation relationship. Appraisal and decision processes are explained and contextualised with aggressive or nonaggressive outcomes (Allen, Anderson & Bushman, 2018). The immediate appraisal is the first process in place. It occurs spontaneously, without conscious control and with no cognitive effort, and is influenced by the present internal state. This process is followed by the decision on how to respond to a situation. This process relies on both time and mental resources, and is influenced by an evaluation of the outcome of the immediate appraisal. If time and cognitive resources are sufficient, and the outcome is evaluated as important, a reappraisal of the event will take place. If not, the person will act according to the behavioural script activated during the immediate appraisal, with very little insight about the nature of the decision made. The reappraisal, when it takes place, can influence internal variables, and it will lead to a more careful evaluation of the situation (Allen, Anderson & Bushman, 2018). It can confirm the immediate appraisal, leading to aggressive thoughts and actions, or it can lead to considering alternatives.

The second main component of the GAM, the distal processes, focuses on the biological and environmental factors that influence the individual. They operate in the background, simultaneously with the proximate processes. Both biological and environmental modifiers "increase the likelihood of developing an aggressive personality" (Allen, Anderson & Bushman, 2018).

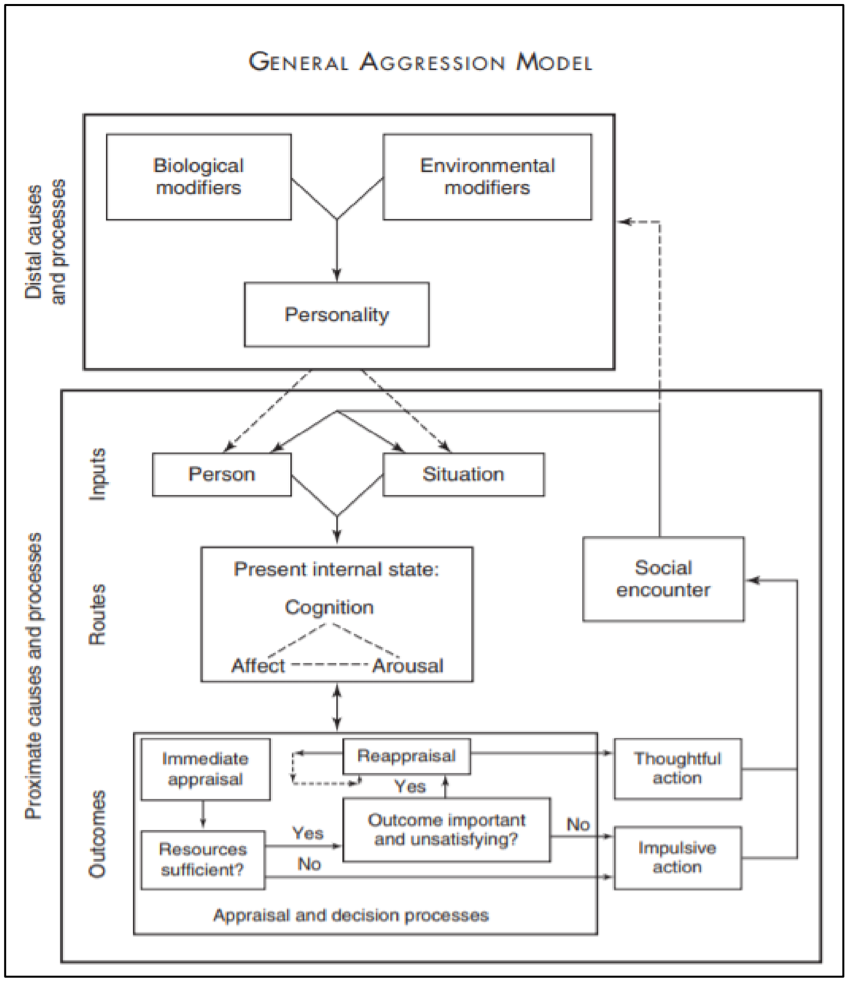


Figure 1. General Aggression Model. From Allen, J. J., & Anderson, C. A. (2017). General Aggression Model. In P. Roessler, C. A. Hoffner, & L. van Zoonen (Eds.) International Encyclopedia of Media Effects. WileyBlackwell.

Figures 2 and 3 give a representation of the relationships among the five theories of aggression that compose the AGM framework.

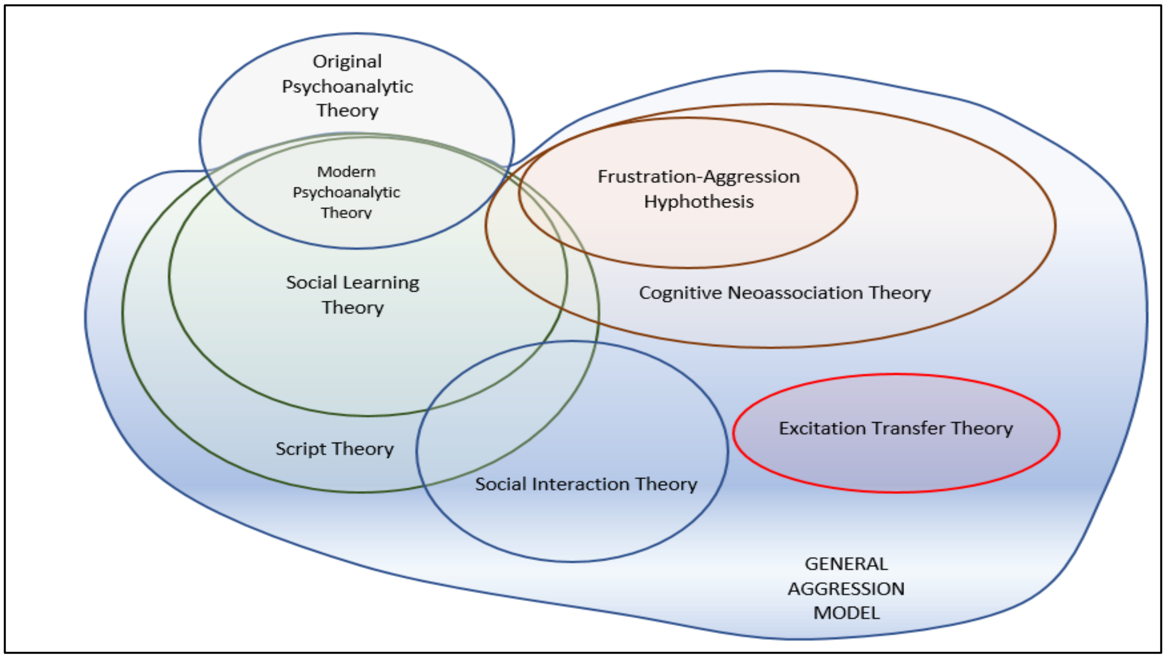


Figure 2. Relationship of AGM to other theories

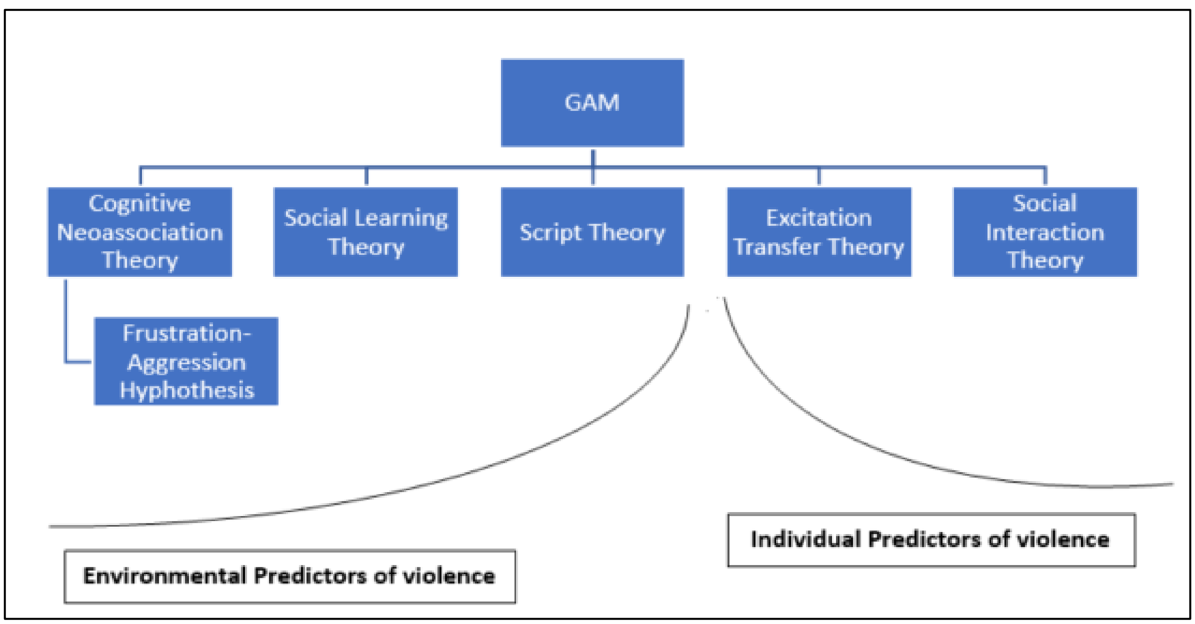


Figure 3. Components of AGM

5. Neurobiology of aggression

Aggression is a complex construct. The neurobiological studies addressing aggression have been mostly focused on understanding the impulsive type, which is the variant typically addressed in the clinical setting, whereas the premeditated variant is typically addressed in the forensic setting. As a consequence, brain regions involved in affect processing, impulse control and emotional decision making have been investigated more extensively than others in this field (Rosell & Siever, 2017).

5.1 Prefrontal cortex, limbic system and prefrontal-amygdala interactions

The role of prefrontal cortex in controlling aggressive and dyssocial behaviours has been initially shown by studying individuals with prefrontal cortical lesions who displayed disinhibited aggressive behaviour (Siever, 2008). The temporal lobe was also identified as a region involved in violent and aggressive behaviours. The most common manifestation is a consequence of tumors and lesions in the temporal lobe, that can result in aggressive behaviours (Tonkonogy & Geller, 1992; Siever, 2008).

Thanks to the adoption of the functional brain imaging, it was possible to make further progresses in this field, investigating patterns of activation in specific brain districts. Several studies have shown anomalous patterns of activation in individuals showing aggressive behaviours. These findings suggest an impaired capacity of the cortical regions to control and suppress the behaviours triggered by provocative stimuli and leading to negative consequences (Siever, 2008). Consequently, impulsive aggression is thought to be caused by a reduced capacity of the individual to control motoric aggressive responses to external stimuli.

Another important finding refers to the role played by the limbic system, that results in hyperactivated impulsive aggression and violence. In 1878, Paul Pierre Broca introduced the concept of the "le grand lobe limbique", or great limbic lobe, using the term "limbic" to describe the arched perimeter of the cortex which includes the cingulate and the parahippocampal gyri (Rajmohan & Mohandas, 2007). In 1937, James Papez postulated a model, know as Papez circuit, assigning to the limbic lobe a role in emotion, and a few years later, in 1952, Paul D. MacLean coined the definition "limbic system" to identify a complex neural substrate for emotion which included Broca's limbic lobe and related subcortical nuclei (Rajmohan & Mohandas, 2007). Researchers have found that in psychopathological contexts there is an alteration of the "top-down" control, operated by specific structures and responsible for calibrating the human behaviour, and an excessive "bottom-up" response elicited by limbic structures, such as the amygdala and insula (Siever, 2008). The structures responsible of the "top-down" regulation are the orbital frontal cortex and anterior cingulate cortex. The "top-down" control systems are responsible for the calibration of the behaviour in response to social cues as well as the prediction of punishment and reward (Blair, 2004). Furthermore, recent studies in the field of neurobiology of aggression have demonstrated the involvement of two limbic prefrontal regions, strictly interconnected with the amygdala: the orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC) (Rosell & Siever, 2017). An altered functioning of the axis OFC-amygdala is thought to negatively affect the capacity to properly attribute affective and motivational significance to stimuli, as well as the "top-down" regulation. An alteration of the axis ACC-amygdala is thought to interfere with both the cognitive modulation of subcortical affect processing and the development of negative self-referential emotional states (Rosell & Siever, 2017).

5.2 Amygdala

The amygdala is a structure located in the medial region of the temporal lobe, and composed of 3 nuclear complexes: basolateral, central or centromedial, and superficial or cortical (Sah et al., 2003; Rosell & Siever, 2017). Its involvement in aggression is well documented. In fact, it is deeply involved in many emotional processes (Salzman & Fusi, 2010), playing a critical and important role in both integrating and transmitting to different cortical and subcortical regions information that is sensory and emotionally relevant for the individual (Rosell & Siever, 2017). Within the general theories of human aggression, the amygdala is "somewhere between cortical decision-making centers and hypothalamic and/or brainstem centers of execution" (Haller, 2018). One consistent finding is the inverse association between the amygdala volume and trait aggression. In addition, many studies begun investigating the functioning of the different subdivisions. Thanks to the recent findings, we know that various subdivisions play different roles (Haller, 2018). Aggression is commonly associated to an increased reactivity of the amygdala to stimuli perceived as socially threatening. However, hyporesponsivity of the amygdala has also been associated with threat, suggesting that some interpersonal/affective dimensions of psychopathy might drive aggressive behaviours (Rosell & Siever, 2017).

5.3 Brain abnormalities

Through brain imaging it was possible to find associations between reduced gray matter and aggression. For example, a growing body of research is establishing the existence of an association between brain abnormalities and antisocial personality disorder and schizophrenia, the two conditions with the greatest implication in violent behaviour (Barkataki et al., 2006). Narayan et al. (2007) found cortical thinning in the medial inferior frontal and lateral sensory motor cortex in violent individuals diagnosed with schizophrenia and/or antisocial personality disorder. Raine et al. (2000) have shown a 11.0% reduction in prefrontal gray matter volume in individuals diagnosed with antisocial personality disorder. Another study found a reduced volume of the whole brain and the temporal lobe in individuals with antisocial personality disorder, while individuals affected by schizophrenia with a history of violence showed a decrease in the whole brain and hippocampal volumes and an increased putamen size, in addition to other abnormalities (Barkataki et al., 2006). A significant volumetric reduction of the anterior cingulate (AC) has also been found in patients with antisocial personality disorder and schizophrenia, with a history of childhood psychosocial deprivation that included physical and sexual abuse (Kumari et al., 2014). A reduced volume of grey matter in the prefrontal cortex associated to an increased volume of grey matter in cerebellar regions and basal ganglia structures was found in a group of 40 male high-risk offenders (Leutgeb et al., 2015).

6. Conclusion

The aim of this chapter is to provide an overview of the main psychological and neurobiological theories associated to violence and aggression. This review will guide us in developing our predictive models in the next chapters. A greater focus on this topic would be necessary for a deeper understanding of the psychological and biological aspects involved in aggression.

Chapter IV

RAI instruments: background, development and adoption

1. Background

In the last decade we assisted to significant changes in healthcare systems, in Canada as well as all over the world. There are several factors affecting service provision in healthcare, such as the demographic changes and the role played by technology in healthcare. An increase in the proportion of the aging population is certainly one of them, and it translates into an increase in the demand for care (Boss et al., 2007). According to a recent report commissioned by the European Commission that focuses on the EU and OECD countries, in 2050 the number of old people aged over 80 will be more than (OECD, 2013). It is estimated that across the OECD countries the share in the population will rise from 3.9% in 2010 to 10% in 2050, while in EU countries it will rise from 4.7% to 11.3%. As mentioned by the report, "between one quarter and one half of them [old people] will need help in their daily lives. Yet governments are struggling to deliver high-quality care to those facing reduced functional and cognitive capabilities (OECD, 2013).

The role played by technology is another key element. The implementation of information and communications technology (ICT) services in healthcare had the purpose of enhancing the quality of care provided to patients. The adoption of information technology poses benefits as well as risks. Certainly, ICT services are critical to make health services more financially sustainable and efficient than paper-based practices. For example, technology automates and extends tasks that were previously accomplished by human personnel. The generation of huge amount of data (patient records) will be managed by replacing paper with digital charts, making patient care more efficient. Genomics and personalized medicine will contribute even more to the generation of huge amounts of information, collecting more patient data and leading to new insights (Thimbleby, 2013).

The quality of available information is extremely important in supporting the transition from a traditional system to a more advanced and efficient model. Policy makers need high-quality data in order to address the complexities emanating from inherent differences at a population and individual levels (Carpenter & Hirdes, 2013).

The challenge posed by the increase of aging population is a factor affecting what funding models need to be adopted. The adoption of a traditional funding system, that uses uniform rates allocated on a per patient basis, has generated a dynamic which penalizes both admission and retention of patients who needed expensive care. As summarized by Hirdes and colleagues (Hirdes et al., 2002), we can affirm that a traditional funding system would penalize a facility that provides more resource-intensive care, whereas a facility with a lighter burden would receive more financial benefits. A more fair and sustainable management of resources in healthcare would require the distribution of funds according to specific patient needs, leading to what many authors call a "case mix-based system" (Hirdes et al., 2002). There are three main types of systems: assessment systems, case mix systems and financing systems. Case mix systems are able to identify groups of patients who require similar resources and categorizing what resources each group needs (Hirdes et al., 1999). Despite "the absolute amount of care provided varies widely, the relative resource needs of different groups of patients tend to be stable across cultures even when the resources available through the financing system vary substantially" (Hirdes et al., 1999). The shift from a traditional system to a more advanced system, such as the aforementioned case mix-based system, is certainly challenging. During the decade of the 80s there was a significant effort in the US to develop better methods of understanding the population of residents in nursing homes. The implementation of the prospective payment system (PPS) for acute care in hospitals has played a key role in bringing the attention to the necessity of extending case mix systems into all the institutional providers. The Centers for Medicare & Medicaid Services (CSM) defines a prospective payment system (PPS) as "a method of reimbursement in which Medicare payment is made based on a predetermined, fixed amount. The payment amount for a particular service is derived based on the classification system of that service" (CSM, 2017). These systems should have the capacity to recognize different care needs of different patient populations, promoting a more equitable allocation of resources (Clauser & Fries, 1992). Strictly related to the case mix systems is the utilization of resident-level assessments, whose information is utilized for different purposes, including case mix measurement. As noted by Clauser and Fries (1992), during the 90s in the US the resident assessment instruments have been considered a key element in the management of clinical care in nursing homes. The government has imposed a process to routinely collect this data, applied to all residents supported by the Medicare and Medicaid programs (Clauser & Fries, 1992).

2. Origin and evolution of the residential assessment instrument

Many authors already recognised the need for a uniform resident assessment in long-term care during the 80s. In 1986, the Institute of Medicine (IOM) has stated that uniform resident assessment as critically important to improve quality and reform the survey process (Morris et al., 1990). In fact, the IOM envisioned a comprehensive assessment of the strengths, preferences and needs of an individual, supported by a customized care plan (Hawes et al., 2007). The following is a brief extract from the report “Improving the Quality of Care in Nursing Homes” issued by the IOM:

> “*Providing high quality of care requires careful assessment of each resident's functional, medical, mental, and psychosocial status upon admission, and reassessment periodically thereafter, with change in status noted... . [The] development of individual plans of care clearly depends on resident assessments*” (Institute of Medicine, 1986, p. 74).

In response to the issues about quality of care in the US nursing homes, the US Congress in the Omnibus Budget Reconciliation Act of 1987 (Public Law 100-203) mandated a nationwide assessment system for nursing home residents with the purpose to provide the needed care plans to nursing home residents. Since 1988, a consortium of different research organizations (Research Triangle Institute in North Carolina, Hebrew Rehabilitation Center for the Aged in Boston, Brown University, and University of Michigan) started working on related initiatives under contract with the Health Care Financing Administration (HCFA). This was the origin of the first interRAI instrument, the Resident Assessment Instrument (RAI) (Clauser & Fries, 1992).

In 1991, most states imposed the implementation of the RAI to nursing homes. In 1992, an international not-for-profit collaboration named InterRAI was founded, bringing together clinicians, researchers and health administrators from over 30 countries. It was the beginning of the ambitious project that led to the creation of first third-generation assessment system (Carpenter and Hirdes, 2013). The vision of this project is that “the assembly of accurate clinical information in a common format within and across services sectors and countries enhances both the well-being of frail persons and the efficient and equitable distribution of resources" (Fries et al., 2003). The RAI was initially designed to "guide individualized resident care planning with two interrelated components" (Clauser & Fries, 1992). The necessity of providing a standardized, comprehensive, and reproducible assessment to each resident was a key point that guided the design of this tool. "With consistent application of item definitions, the RAI ensures standardized communication both within the facility and between facilities. Basically, when everyone is speaking the same language, the opportunity for misunderstanding or error is diminished considerably" (U.S. Department of Health and Human Services - Health Care Financing Administration, 1995). The RAI is composed of three main components: the Minimum Data Set (MDS), Triggers and Residents Assessment Protocols (RAPs), and Utilization Guidelines. It is intended to be completed by an interdisciplinary clinical staff, including nurses, social workers, occupational therapists, pharmacists, and the attending physician who can provide important information to the MDS and the RAPs. According to the federal regulation, each individual who completed a portion of the RAI was required to sign, date, and certify its accuracy. In addition, a registered nurse was required to sign and certify the completion of the assessment (Department of Health and Human Services, Office of Inspector General, 2001). The HCFA began developing the version 2.0 in early 1993.

2.1 RAI components

I. The *Minimum Data Set (MDS)* is a component developed to contain a group of core items to assess nursing home residents. With more than 300 individual items, the MDS can be defined as the heart of the RAI. It was the result of an extensive effort to incorporate data to measure residents' strengths as well as psychosocial needs, validated by solid testing and enriched thanks to the collaboration of hundreds of clinicians, health care professionals and regulators. (Clauser & Fries, 1992). "The items in the MDS standardize communication about resident problems and conditions within facilities, between facilities, and between facilities and outside agencies" (U.S. Department of Health and Human Services - Health Care Financing Administration, 1995).

II. The *Resident Assessment Protocols (RAPs)* are a component of the utilization guidelines generated by single items or a combination of MDS elements that are "triggered" (selected). They represent "structured, problem-oriented frameworks for organizing MDS information, and examining additional clinically relevant information about an individual" (U.S. Department of Health and Human Services - Health Care Financing Administration, 1995). RAPs address 18 major domains in nursing home, providing guidelines for the development of care plans.

III. The *Utilization Guidelines* give instructions about when and how to use the RAI.

2.2 Case mix systems: RUGs

The development and application of case mix resident classification systems for nursing homes in the US was possible thanks to the contributions of the HCFA that founded different initiatives. These systems are called Resources Utilization Groups (RUGs). Their main goal is to understand and explain how resources are utilized. This goal is achieved by grouping home residents according to residents' characteristics (Clauser & Fries, 1992). RUGs have been introduced for the first time in 1985 in the US. Since then, they have been refined several times, evolving through three versions: RUG-I, 8 categories; RUG-II, 16 categories; and RUG-III, 44 categories. The third version (RUG-III) has been developed to improve the discrimination capacity to categorize low-volume/high-cost/high-acuity patients by considering medical conditions, treatments/services and psycho-social factors, in addition to Activities of Daily Living (ADLs) and Behaviours of Daily Living (BDLs) (Botz et al., 1993). The groups identified by the RUGs represent mutually exclusive categories, defined by the resource utilization in long-term care settings. There is an important relationship between the RAI and the RUGs. In fact, the Minimum Data Set of the RAI previously described (RAI/MDS) provides the clinical data based on which the RUG-III system has been built: the RAI/MDS is used to assign individuals to each category (Tucker, 2009). Once the items are triggered, the responses are used to calculate the RUG-III. Section G of the RAI/MDS, which refers to ADLs and mobility assessments, provides a significant contribution to the generation of the RUG-III classifications (Ellen Dellefield, 2006). The RUG-III can classify a resident in 44 categories, organized into three dimensions. The first dimension is represented by seven major clinical domains: rehabilitation, extensive care, special care, clinically complex, cognitively impaired, behavior problems, and reduced physical function. The second dimension is related to Activities of Daily Living (ADLs). The third dimension addresses the rehabilitation services, indicating also the possible presence of depression (Chou, Chi, & Leung, 2010). The residents' classification process can be divided in two steps: as a first step, residents are assigned to one of seven major categories; successively residents are classified into 1 of 44 minor categories. In 2006 there was an expansion of the RUG categories, from 44 to 53. The implementation of the third version of the Minimum Data Set (MDS 3.0) was scheduled for October 2010, and the utilization of the corresponding RUG, the RUG-IV, was scheduled for October 2011, as decided by the Congress during the passage of the Patient Protection and Affordable Act (Tomaino, 2010).

The original RAI has guided the development of all the assessment instruments, collected as an integrated family of assessment instruments (www.interrai.org), realized at a later time.

Their design was thought to utilize all the information available, with the intent to serve multiple audiences including care panning, outcome measurement, quality improvement, and resource allocation. The creation of the first RAI instruments, realized following a serial process, acted as a guide for the subsequent instruments. Subsequently, the initial instruments have been progressively updated in parallel. It became clear that the entire set of instruments had to be considered as an integrated system (Hirdes et al., 2008). Other interRAI instruments developed in the 1990's include:

- *The RAI for Acute Care* (RAI-AC), realized with the partnership of different countries including Canada, US, UK, Norway and Iceland, and finalized in the fall of 1998 (Hirdes et al., 1999). It is a multidimensional geriatric assessment providing a holistic picture of hospitalized older persons, able to determine their medical, psychosocial and functional capacity and needs (Wellens et al., 2010; Devriendt et al., 2013).
- The *RAI 2.0* is a system developed for nursing homes, specifically designed for complex and frail seniors. It has been mandated for use in Ontario chronic care hospitals and long-term care facilities in 1996 (Tjam et al., 2012; Hirdes et al., 1999). The core of the RAI 2.0 is the Minimum Data Set (MDS) 2.0, composed by over 300 individual items, and Clinical Assessment Protocols (CAPs). This instrument was developed to address a wide range of domains. For this purpose, several outcome measures have been developed and validated, such as scales to assess cognition function, behaviour, depression, social engagement, etc. (Tjam et al., 2012).
- The *RAI for Home Care* (RAI-HC) is a comprehensive assessment instrument to support home health and home care services. I was originally developed in 1994, and five years later Version 2.0 was released (Hawes et al., 2007). This instrument addresses issues

related to the functioning and quality of care in adults in home care settings, frequently used with both frail elders and individuals with disabilities (Hawes et al., 2007). The RAI-HC consists of two parts. The first part is designed to collect information from a broad spectrum of domains. The RAI-HC's items share 47% of those in the RAI 2.0 (Morris et al., 1997). The second part of this instrument is composed of 30 problem focused clinical assessment protocols (CAPs) covering various conditions that represent common risks in home care (Hawes et al., 2007).

- The *RAI for Post Acute Care* (RAI-PAC) was developed to target elderly patients who need rehabilitation as well as short-stay clinically complex patients (Hirdes et al., 1999). This tool was designed to collect "information on sociodemographic, clinical, and functional indicators of health status in patients admitted to PAC, and is one of a suite of instruments used worldwide in nursing homes, mental health facilities, home care, and acute-care setting" (Gindin et al., 2006). The updated version of this instrument is the interRAI *Post-Acute Care and Rehabilitation (PAC-Rehab)* Assessment, and it can be used as a valuable ally of either the interRAI Acute Care (AC) Assessment or the interRAI Acute Care for Comprehensive Geriatric Assessment (AC-CGA) Systems (http://www.interrai.org/).
- The *RAI for Mental Health* (RAI-MH) is a comprehensive assessment instrument for psychiatry designed to support a variety of applications and satisfy the needs of different audiences (Hirdes et al., 2001). It is the result of the effort of a six-country research team (Canada, the United States, the United Kingdom, Japan, the Netherlands, and Norway), under the guidance of the Ontario Joint Policy and Planning Committee's (JPPC) Psychiatric Working Group (Hirdes et al., 1999). The previous version, the RAI-MH 2.0

has been replaced by the interRAI-MH, and it is being implemented in both Iceland and Finland (http://www.interrai.org/)

The interRAI instruments are the answer to the necessity of having highly specific data that a facility must provide to the government, in order to describe with precision and reliability their patient population and the services provided to them. With their development, the standardized assessment instruments are capable of providing complex and highly detailed information, to assist and support the government in the delicate process of resource allocation.

2.3 Quality indicators (QIs)

A key element of the interRAI instruments is a group of indicators developed to monitor and assess the quality of care. The scores reported by these indicators, in turn, are used to evaluate the effectiveness of the resources allocation. A strong contribution to the development of the QIs has been given by the nursing home sector during the 1990s. The QIs operate as a “foundation for both external and internal quality-assurance (QA) and quality-improvement activities (Zimmerman et al., 1995). produced significant effort has been done by the Center for Health Systems Research and Analysis (CHSRA), at the University of Wisconsin–Madison, where a group of researchers have utilized data coming from the Resident Assessment Instrument (RAI) to develop a set of indicators to monitor the quality of care. The creation of this set of QIs follows some previous developments in nursing homes, including the growing interest in understanding the issues connected to both quality of life and quality of care in nursing homes (Zimmerman et al., 1995). Thus, having a set of QIs is a key point to evaluate the performance of a facility in delivering care. They are one of the most important tools to make decisions and set priorities. For example, they can inform about the cost-effectiveness of a specific treatment or care plan.

2.4 Evolution of the assessment instruments

The role played by the information itself is certainly central in this discussion (Hirdes 1999 et al., 1999). Significant benefits deriving from the implementation of a standardized assessment tool refers to the capacity of integrating the information across different domains in healthcare, with advantages for patients, providers and governmental agencies (Hirdes et al., 2002). In fact, a patient who receives an assessment does not necessarily have to go through the process multiple times. It is a strategy to guarantee the continuity of care, reducing at the same time the assessment burden. On the other hand, a high degree of integration of the information would improve the quality of the communication among different organizations and facilities (Hirdes et al., 2002). For example, a standardized assessment tool would allow a mental health organization to collect data that can be successively integrated with other sectors. The implementation of a standardized tool would successfully address the issue of the lack of standardized data deriving from assessments developed in-house by mental health providers. The standardized assessment instruments can be classified in three generations (Carpenter and Hirdes, 2013).

- The first generation refers to mono-dimensional tools, composed by scales that are build to "measure a single construct for a single purpose" (Carpenter and Hirdes, 2013). Some examples are the Mini Mental State Examination – MMSE (Folstein et al., 1975), Geriatric Depression Scale – GDS (Yesavage et al., 1982).
- The second generation includes multidimensional instruments that address multiple domains. These instruments are more complex than the previous ones and are developed to play a key role in supporting care planning. In addition, the

data can be used to generate different measures of outcome, case mix, quality of care and eligibility criteria (Hirdes et al., 1999; Carpenter and Hirdes, 2013).

- The third generation of instruments represents a progress of the second-generation tools. They are a set of longitudinal and person-focussed instruments, specifically designed to be used by healthcare professionals, finding application with patient populations and care settings. In that regard, in each interRAI instrument we find a set of items that are present in every care setting, as they address key aspects of the principal domains in the individual. In addition, there is a set of data that are specific to each care setting. The recorded information of an individual might address multiple domains and stages of life, for example describing changes in their abilities, or addressing also multiples care settings (such as residential care, mental healthcare, community care).

The implementation of protocols shared by various settings and the capacity to efficiently move the information stored across different settings make them an excellent strategy to support continuity of care. The interRAI suite is currently the first third-generation assessment system (Carpenter and Hirdes, 2013). An important asset of these tools lies in the possibility to use them regardless of the care setting. One element that is critical is the presence of various Clinical Assessment Protocols (CAPs), supported by multiple clinical scales and algorithms, necessary to understand and interpret the clinical findings. The interRAI instruments have been developed with a uniform structure. The backbone of each interRAI tool is the data set, specific to each care populations and care settings. A set of core items, utilized for all care populations and care settings, is supported by an addition set of items that are specific to the population and setting to which the instrument refers (e.g. mental health, home care etc.). Each interRAI instrument has incorporated the following elements (see Figure 4): a set of assessment items, a set of scales, a set of algorithms, quality indicators and case mix measures to resource use (Carpenter & Hirdes, 2013):

I. The assessment items

The assessment items are organized in sections, each one addressing a specific area such as health conditions, physical condition, and cognition. All the interRAI instruments have a minimum number of data fields to record patients' data, called Minimum Data Set (MDS), that include time frames.

II. The Clinical Algorithms - Clinical Assessment Protocols (CAPs)

The Clinical Assessment Protocols (CAPs) are clinical algorithms, generated to support the utilization of care plans. Each CAP refers to a specific domain, created by an international consortium of experts (clinicians and researchers). They play an important role in the interpretation of the collected data thanks to the Minimum Data Set (MDS). The clinical algorithms are connected to guidelines addressing specific problems referred to health, psychological, social and environmental domains (Hirdes et al., 1999). The CAPs provide all the necessary information, so that they can be used as a sort of clinical guideline. They describe the condition or issue targeted, explaining what items are involved in the generation of the score, and identifying the recommended goals of care (Carpenter and Hirdes, 2013). The clinical algorithms utilize a software to generate a set of scales once all the sections of an interRAI have been completed. These scales can serve the purpose of diagnostic screening or can provide severity measures (Carpenter and Hirdes, 2013).

III. Performance scales

The algorithms generate performance scales related the individual being tracked. The information collected can refer to self-reported responses to questionnaires or to a comparison between the observed performance and the items triggered (Carpenter and Hirdes, 2013). Some examples are the Cognitive Performance Scales, the Depression Rating Scale, RAI/MDS Health Status Index (measuring the quality of life), the three Activity of Daily Living Scales, and the Index of Social Engagement (Hirdes et al., 1999).

IV. The quality indicators (QIs)

The quality indicators are indicators of process and outcomes of care. They can be used not only on a national level but also internationally, to compare how different countries perform in specific care settings. In that regard, a pilot study has compared the performance in home care services in 11 European countries (Czech Republic, Denmark, Finland, France, Germany, Great Britain, Iceland, Italy, the Netherlands, Norway and Sweden), focusing on a population of 4,007 individuals aged 65 and over. (Bos et al., 2007; Carpenter and Hirdes, 2013). The indicators used in this case were the Home Care Quality Indicators (HCQIs), derived from the Minimum Data Set specifically for Home Care. The study revealed that Czech Republic, Italy and Germany had the worst outcomes, and they need a closer examination (Bos et al., 2007).

V. The case mix measures to resource use

The interRAI tools categorize residents into different clinical categories, identified by the case mix systems, reflecting the costs associated to services utilized and supports needed. These algorithms "provide a person-specific means of allocating health care resources based on the variable costs of caring for individuals with different needs" (interRAI, 2018). Case mix systems are very efficient in exploring both the relationships between costs and patient needs, and costs and quality of care (Carpenter et al., 1997). In many countries RUGs is used as a basis of payment for funding long-term care. They use a group of items that belong to the Minimum Data Set (MDS) to calculate the cost of caring for a nursing home resident (Carpenter and Hirdes, 2013). The third version of the RUG system (RUG-III) has proven to be useful in a variety of settings (Carpenter et al., 1997).

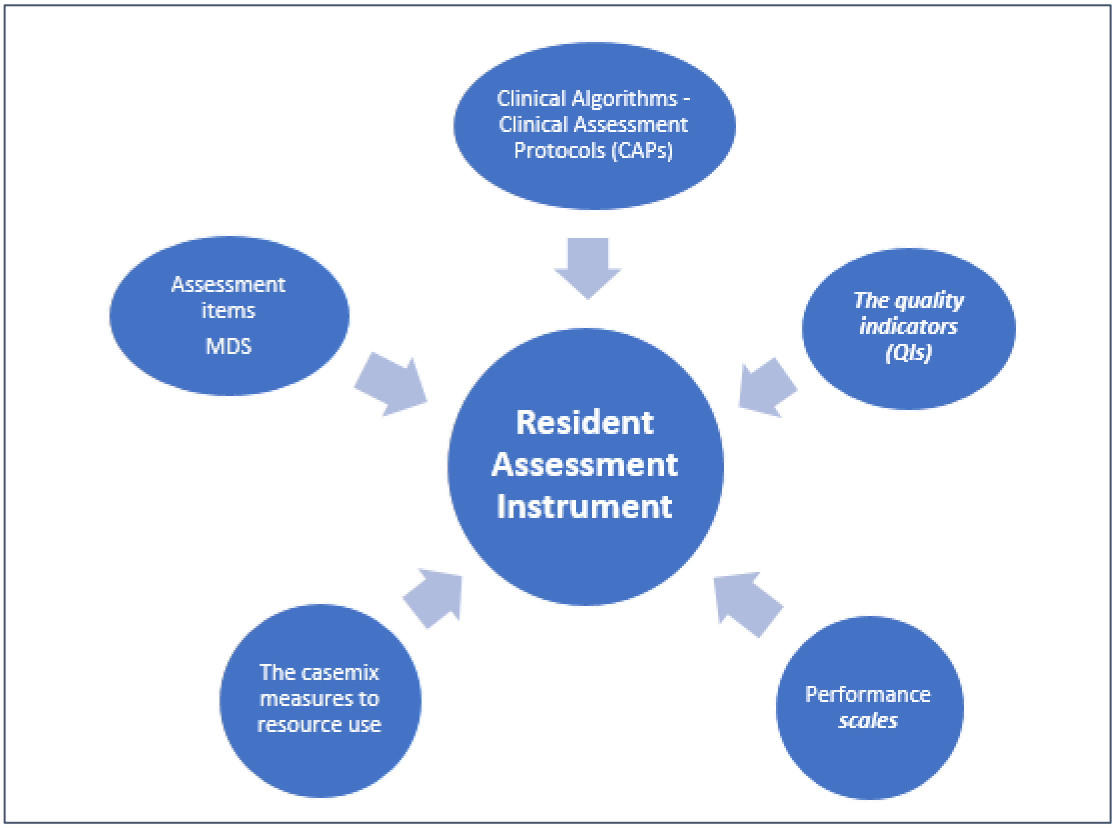


Figure 4. Components of the RAI

The implementation of advanced third-generation assessment instruments poses significant challenges. First of all, a political commitment is required. This is the first step towards a deep transformational change. Financial commitments are also necessary. Given the technology requirements needed to use the interRAI instruments, it is important that facilities make investments to have appropriate IT infrastructures. This appears to be the successful route to achieve high quality care and performance data in health care. Education and training are also critical aspects. A cultural change is required within an organization at various levels, from the top management to the clinical staff in charge of directly using these instruments. An effort should be made to integrate the assessment systems into routine clinical and social care practice. Finally, given their experimental nature, these tools need to be assessed and continuously improved to better serve patients and the clinical staff.

3. Resident Assessment Instrument-Mental Health (RAI-MH)

The challenges faced by the mental health care led to reconsider the funding systems for psychiatry. In the previous chapter, we have discussed the pathway that brought other health care sectors, such as the nursing home sector, towards the implementation of case mix systems, to guarantee a more equitable distribution of payments.

In the 1990s the Ontario Joint Policy and Planning Committee (JPCC) begun a collaboration with the Clarke Institute of Psychiatry Consulting Group, with the purpose of investigating possible solutions to implement a funding system for psychiatric hospitals based on Psychiatric Case Mix Groups (CMGs), developed by the Canadian Institute for Health Information (CIHI). As a result, the Psychiatric Working Group was formed (PWG) and it comprised various mental health stakeholders. The original mandate of the PWG was to develop a patient classification system for hospital-based psychiatry

(Hirdes et al., 2000). After the PWG was formed, the JPPC's Chronic Care and Rehabilitation Working Group recommended the utilization of the Resident Assessment Instrument 2.0 (RAI 2.0) for chronic care hospitals in Ontario, an instrument designed for long-term care use. The question whether or not it was necessary to have a similar tool specifically for mental health was just a matter of time. Several international researchers, including the staff involved in the interRAI project, had the opportunity to meet and present their work related on case mix and assessment in mental health. Different discussions that took place between experts, supported by an extensive review of the available literature, led to a new effort to develop the RAI-MH. One critical finding that emerged during these discussions was the necessity to create a strategic bridge between the information about process and outcomes of care and case mix algorithms. Another finding was the necessity to establish a benchmarking system with the capacity of addressing some potentially negative outcomes or process of care in the case mix algorithm (Hirdes et al., 2000). A study of 2,300 psychiatric patients found that a per diem case mix model explained approximately 33% of resource utilization (Hirdes et al., 2000; Yamauchi, 1997). It became clear that the mental health care would have benefited from a funding system more equitable, and a resident assessment instrument specifically developed for mental health would have been the first step towards this goal. This instrument would be crucial for address care planning, quality improvement, outcome measurement, and case mix (Hirdes et al., 2002).

It was decided that the new instrument should serve all the inpatient psychiatric adults, including acute and forensic psychiatry. The realization of the RAI-MH project has involved clinicians and researchers from several countries: Canada, the United States, the United Kingdom, The Netherlands, Norway, and Japan. The leading role was under the responsibility of a Canadian research team based in Ontario, supported by the JPPC Psychiatric Working Group (PWG) (Hirdes et al., 2002). Thus, the RAI-MH developed to be an instrument with the same characteristics of the other RAI instruments: trigger items associated with algorithms, scales, quality indicators and case mix measures to resource use. In line with the utilization purpose of the other interRAI tools, the intent of the RAI-MH is to support clinicians in making decisions. A key element of the RAI-MH was the development of a series of protocols, originally called the Mental Health Assessment Protocols (MHAPs). Each one of these protocols uses a defined set of trigger items, to flag a current or imminent problem that clinicians need to address. Thanks to these clinical triggers it is possible to identify individuals who need further support in specific domains (Martin et al., 2009). MHAPs are a critical feature of the RAI-MH. They are intended to give a multidimensional perspective of the resident to the clinical staff, and can be used to support the creation of individualized care plans. It is important to underline that the MHAPs are not diagnostic tools; rather they are a strategic ally that helps the process of assisting clinicians in the identification of issues emerging in the daily life and areas of improvement (Martin et al., 2009). Some examples of MHAPs are: Violence, Self-Harm, and Abuse by Others. These protocols provide a description of the clinical problem, describe the algorithms involved in the identification of the problem, provide a brief literature review, and a set of recommendations to address the problem (Hirdes et al., 2002). The MHAPs have been renamed Clinical Assessment Protocols (CAPs) (Hirdes et al., 2008).

Similar to other interRAI instruments, the RAI-MH is equipped with a set of quality indicators. Initially a group of 32 quality indicators, defined as Quality Indicators for Mental Health (QIMHs), was developed to measure both process and outcomes of care (Hirdes et al., 2000). This group was further refined and expanded to 35 indicators (Perlman et al., 2013). Process indicators focus on aspects such as quality, efficiency, safety, or accessibility of services delivered. Outcome indicators are particularly useful to evaluate the effectiveness of mental health services (Perlman et al., 2013). When these two types of indicators are used in combination, they can provide further insights.

The RAI-MH can find successful applications to measure different types of outcomes. In that regard, there 10 outcome scales embedded in the current version of the instrument, many of which were adapted from the previous versions of the RAI-MH (Urbanoski et al., 2012). Some examples are:

- the Cognitive Performance Scale (CPS), designed to utilize data from the MDS and provide information about the cognitive performance of residents (Morris et al., 1994).
- The Minimum Data Set Depression Rating Scale, particularly useful to conduct screening for depression in nursing-home residents (Burrows et al., 2000)

- A pain scale, developed from the MDS, to assess pain in nursing home residents (Fries et al., 2001).
- A scale to measure activities of daily living (ADLs), based on MDS data, providing useful information to evaluate a resident's ADL status (Morris, Fries and Morris, 1999).

Two other scales that were embedded in the RAI-MH are worthy of mention; the GAF (DSM-IV) and the CAGE-AID (Erwing, 1984; Brown & Rounds, 1995):

- The Global Assessment of Functioning (GAF) assesses the psychological, social and occupational impairments of an individual, excluding impairments related to physical or environmental limitations (DSM-IV). A clinical judgement is assigned through a number that ranges from 0 to 100 (DSM-IV).
- The CAGE-AID (Cut down, Annoyed, Guilty, Eye-opener–Adapted to Include Drugs) is a conjoint questionnaire that expands the previously developed CAGE questionnaire, composed of four clinical interview questions to diagnose alcoholism, into a version that includes alcohol and other drugs (Brown & Rounds, 1995).

A significant effort was made to evaluate the psychometric properties of this instrument, especially with cross-national comparisons. Inter-rater reliability, internal consistency and convergent validity of the items and outcome scales embedded in the RAI-MH have been tested previously, showing solid results (Martin et al., 2009; Martin

& Hirdes, 2009). Different studies have shown acceptable to good reliability of both a significant number of the items in the RAI-MH and summary measures (Urbanoski et al., 2012; Martin et al., 2009; Hirdes et al., 2002; Hirdes et al., 2008). Hirdes et al. (2002) reported the first set of evidence on the reliability and validity of the RAI-MH. Their study has analyzed the bulk of items utilized in the final Version 1.0 of the RAI-MH.

In Ontario the adoption of the RAI-MH as the assessment platform for adults in inpatient psychiatry settings started in 2005 (Urbanoski et al., 2012; Hirdes et al 2000; Martin & Hirdes, 2009). The provincial mandate is to complete the RAI-MH within 72 hours of admission, every three months (quarterly), when there are changes in clinical status, and upon discharge. The current version of this instrument counts around 400 items. Its completion takes from 60 to 75 minutes (Urbanoski et al., 2012; Martin & Hirdes, 2009; Hirdes et al., 2000; Hirdes et al., 2002). It is used by a variety of users (clinicians, researchers, frontline workers, policy-makers, etc.). Designed to decrease significantly the assessment burden, this is a comprehensive, person-centered assessment particularly important to support clinical decision-making (Martin & Hirdes, 2009). Martin et al. (2013) have argued that the RAI-MH can be an important ally in providing the necessary information to facilitate shared-decision making in inpatient mental health settings, educating and empowering individuals with mental illness. Patients have appreciated receiving a narrative summary of their symptoms, the diagnosis, the medications administered, and the strengths. it’s the RAI-MH’s richness of information can serve multiple purpose, including the possibility to educate and empower individuals by using clinical information and support psychiatric rehabilitation initiatives. Urbanoski et al. (2012) conducted a performance evaluation of the RAI-MH at a large psychiatric hospital in Ontario, considering a window of time of the 3 years, from 2005 to 2007. Their analysis has revealed positive and negative results: data quality improved over time, but a greater effort is required to improve the outcome monitoring capacities of the RAI-MH (Urbanoski et al., 2012). Their results suggest the importance of finding solutions for streamlining the assessment and increasing staff compliance, in order to improve the global performance of the RAI-MH.

Chapter V

Statistical prediction methods in violence risk assessment

1. Introduction

The concept of individuality refers to the combination of multiple characteristics that describe each individual (e.g., sex, age, physical condition, medical condition, social-status, etc.). This complexity brings challenges to research in health care, as it is very common to obtain heterogeneous research groups, even if the individuals are grouped by the same condition (a specific disease, or any other health related condition). Thus, pursuing homogeneity, in terms of measurement, by operating a further division of these heterogeneous groups into homogeneous subgroups is a desirable step, from both a clinical and operational point of view (Harper, 2005). In medical decision making, homogeneity brings some benefits, such as enhancing certainty in clinical diagnosis. The classification method utilized should be able to identify precise classifiers, and at the same time supporting the comprehension of the predictive structure of the data (Harper, 2005).

The primary goal in risk prediction is the development of an instrument having a high predictive capacity (Hamilton et al., 2014). In criminal justice, for example, risk assessment instruments are designed to predict behaviour, resulting in modestly accurate predictions (Gottfredson & Moriarty, 2006). On the other hand, if risk assessments are used as a management tool in non-predictive situations they show a low predictive capacity, especially when these instruments are not correctly used (Gottfredson & Moriarty, 2006).

In general, the risk assessments used in criminal justice can be grouped in two main categories; tools based on clinical judgement, and tools based on actuarial practice (S. D. Gottfredson & Gottfredson, 1986; Gottfredson & Moriarty, 2006):

- The first generation of tools for risk assessment is represented mainly by unstructured professional judgement, with a variation called "structured clinical judgement" (Andrews, Bonta & Wormith, 2006). They are also called subjective assessments or professional judgements, and refer mainly to a set of non-standardized questions used to assess the offenders.
- The second generation of tools, also known as actuarial assessments, comprises a set of empirically based tools composed mainly by static items, without the support of a theoretical framework on the background (e.g. Violence Risk Assessment Guide [VRAG]; Harris, Rice, & Quinsey, 1993; General Statistical Information for Recidivism [GSIR]; Bonta, Harman, Hann, & Cormier, 1996). Despite a good predictive validity of some second-generation instruments, the main critique referred to the incapacity of static factors to address the complexity of human behaviour and functioning (Yang, Wong & Coid, 2010).

- The third generation instruments, also referred as "risk-need" instruments, are empirically based, composed by a broader set of dynamic items compared to the previous generation. Some of these methods were built with a theoretical foundation (Historical, Clinical, and Risk Management Violence Risk Assessment Scheme [HCR-20]; Webster et al. 1997; the Violence Risk Scale [VRS]; Wong & Gordon, 2006; Level of Service Inventory and revised version [LSI/LSI-R]; Andrews & Bonta, 1995).
- The fourth generation of instruments combine systematic intervention and monitoring with the assessment of a wider spectrum of risk factors and other relevant personal factors (Andrews, Bonta & Wormith, 2006).
- To overcome some of the limitations posed by the risk assessment tools, a possible fifth generation instrument has been proposed. Unlike previous methods, it is based on non-regression methodologies (Shaffer et al. 2011; Hamilton et al., 2014).

Violence prediction poses several problems, especially when we want to focus our attention on low or very low frequency events, such as trying to predict the executor of the next school shooting (Yang et al., 2010). In these cases, one of the most common problems is related to the risk of obtaining false positive error rates, mistakenly identifying some individuals as violent. On the other hand, the development of actuarial instruments described above, whose adoption became a standard practice in forensic risk assessment, has given a strong contribution to the identification of valid predictors of violent behaviours (Yang et al., 2010).

In the remainder of this chapter, we review three most commonly used statistical prediction methods in violence risk assessment. Many statistical methods have been adopted to address risk assessment prediction, including multiple regression, clustering approaches, and neural network models. Different reasons may lead to prefer a certain method, with different outcomes related to the results obtained (Gottfredson & Moriarty, 2006).

2. Logistic Regression

The logistic regression is a statistical technique applied in many fields of psychology and psychiatry (Liu et al., 2011). It is used when there are one or more independent variables determining an outcome, represented by the dependent variable. To estimate the maximum probability that an event occurs (e.g., a violent act), the dependent variable is initially transformed into a logit variable (the natural log of the odds of the variable). After this transformation, the dependent variable becomes a dichotomous variable with its value ranging from 0 to 1. Thus, this technique provides an estimation related to the occurrence of a certain event. The logistic regression does not share some of the main assumptions that we find in linear regression and general linear models. Firstly, it does not assume a linear relationship between the dependant and independent variables. Secondly, the normal distribution of variables is not required. Thirdly, the homoscedasticity is not needed. Finally, the independent variables can be nominal as well as ordinal.

Amore et al. (2008) used the logistic regression to investigate what factors were independently associated with physical aggression among psychiatric patients one month before the admission and during hospitalization. Hartvig et al. (2009) used the logistic regression to analyse a checklist containing 33 items, the PS (Preliminary Scale), developed to assess violence risk among patients discharged from acute psychiatric facilities. This statistical method was used to determine the odds ratios (ORs) of all the items, and multivariate logistic regression was "performed to test each item's adjusted significance" (Hartvig et al., 2009). Thomas et al. (2005) conducted a comparison of different statistical modeling techniques to predict violence in a sample of patients with psychotic illness. The authors compared the predictive performance of logistic regression and classification tree methods, and concluded that "the full logistic regression had the best overall performance taking into account both sensitivity and specificity" (Thomas et al., 2005). However, Steadman et al. (2000) argued that an approach based on the use of classification trees in violence risk assessment should be preferred to the logistic regression model. They argued that a classification tree approach would be able to reflect the "real-life clinical thinking" about the complexity of the nature of violence better than any other actuarial method. One major critique of the logistic regression is the fact that it ignores that violence in different subgroups can be predicted by different variables (Liu et al., 2011).

3. Classification Trees and Random Forests

The classification tree (CT) model is based on a hierarchy of questions to be answered, and a final decision to be made on the basis of the previous answers (Breiman et al., 1984). Decision trees are “non-parametric question-decision models” (Hamilton et al., 2014) that categorize the data using a set of conditioning answers (Hamilton et al., 2014; Liu et al., 2011). The questions to be answered depend on the answers given to the previous questions. The first questions is posed to all the individuals. Then, depending on each answer (or the nature of the question), a second question is selected and posed. This process continues until each individual is classified. In the field of violence risk assessment, a classification tree approach would be able to reflect a model of violence that is interactive and contingent, improving the classification of an individual as high or low risk by enhancing the combinations of risk factor (Steadman et al., 2000). In other words, the data is split in groups, through a repetitive process, each time using the best possible predictor. Different classification tree models have been developed, finding application in various health care fields, including recent applications in violence risk prediction (Liu et al., 2011). The Classification and Regression Trees (CART) was successfully applied in creating case-group mix, minimum data set requirements, intensive care and hospital inpatients (Harper, 2005). Some types of classification tree models include the Classification and Regression Trees (CART or CRT) (Breiman et al. 1984), Decision Tree Forests (Breiman 2001), Boosting Trees (Friedman 1999a, b), and the Iterative Classification Tree (Steadman et al., 2000). In the field of violence risk assessment, some authors have compared regression models with the classification tree models. Lui et al. (2011) have summarized these findings, reporting that only two studies out of eleven showed similar accuracy; one study reported no difference between the two approaches, whereas another study reported a better performance of the regression model; finally, seven studies reported an enhanced performance of the iterative classification tree models, without supporting their conclusion adequately. The repetitive process of classification trees generates groups with higher degree of homogeneity, a desirable characteristic in health care settings as it increases certainty in individual patient needs and resources utilization (Harper, 2005; Liu et al., 2011).

Random forests (RF) are a combination of multiple trees, generated by the bootstrapping technique from the same data; with random forests it is possible to identify an “average” tree that will be used for predictions (Hamilton et al., 2014). Despite the lack of studies that formally compare the predictive performance of risk assessment between regression models and decision trees and random forests (Hamilton et al., 2014), random forests show interesting qualities. They do not overfit (meaning that a model does not lose predictive accuracy when applied to validation samples), and selecting the appropriate degree of randomness enhances their capacity to act as precise classifiers and regressors (Breiman, 2001). By using a large number of trees, random forests can inductively identify interaction effects between the predictors and the response, and thus allowing a relative cost of both forecasted false positives and false negatives (Berk et al., 2009; Hamilton et al., 2014)

4. Neural Networks Models

Neural networks (NNs) emerged from research in artificial intelligence. They are artificial models attempting to mimic the functioning of human brain (Harper, 2005). A neural network is composed of several interconnected units, known as *neurons* or *nodes*, able to communicate to each other following specific patterns. The functioning of each neuron is limited to a restricted rule, using a process input signal in order to calculate an output signal. The outputs signals are sent to other units using specific connections known as *weights*, which excite or inhibit the signal (Harper, 2005). Multiple neuron layers, that include hidden, or intermediary layers, are used for modelling more complex and nonlinear relationships. Through a connectionist approach to computation, a neural network changes its own structure while learning. This characteristic makes them efficient in modeling complex relationship between inputs and outputs and identifying complex patterns not identifiable by the human brain (Hamilton et al., 2014; Liu et al., 2011).

To provide an example, Figure 1 illustrates a common type of a neural network with one input layer, one hidden layer and one output.

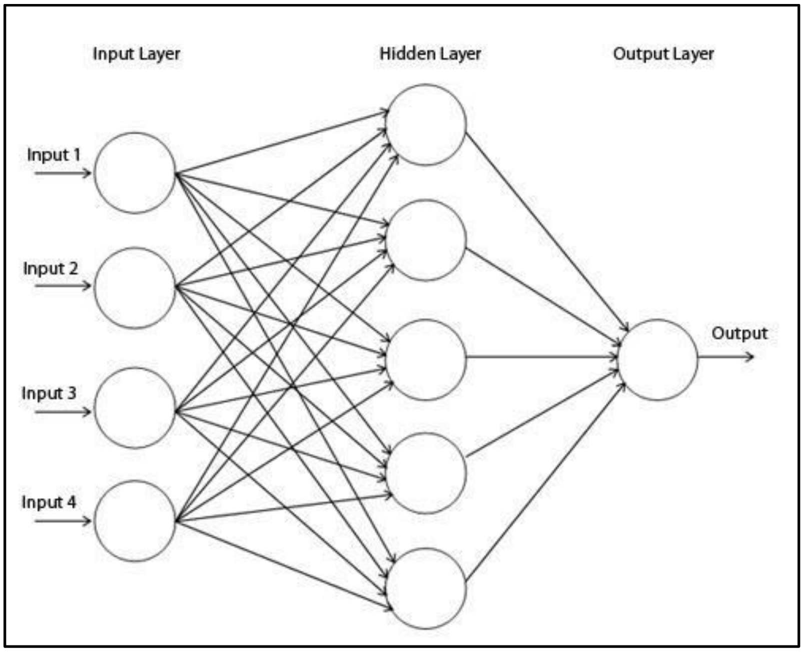


Figure 5. An example of a neural network (source:[https://stackoverflow.com/questions/33649645/how-should-nodes-be-connected-in-a-neural-network on 08/03/2018](https://stackoverflow.com/questions/33649645/how-should-nodes-be-connected-in-a-neural-network).) Retrieved February 6th 2018

The information is passed from the four (4) input units to the five (5) hidden units. The hidden units sum the input units applying specific weights (w_{ih}) adding a constant (named the bias), and take a specific activation function φ_h of the result. Activation functions are very important as they introduce non-linear properties to the network, converting an input signal or a set of input signals of a node to an output signal. The output unit has the activation function φ_0 .

$$f(x) = \varphi_0\left(w_0 + \sum_{h=1}^{5} w_h * \varphi_h\left(w_{0h} + \sum_{i=1}^{4} w_{ih}x_i\right)\right.$$

The weights are the parameters of the mode, and their number is equal to the sum of arcs connecting the units and bias terms (Hamilton et al., 2014). The activation function of the hidden layer is frequently the logistic function (Hamilton et al., 2014):

$$\varphi(z) = \frac{\exp(z)}{1 + \exp(z)}$$

The application of neural networks in risk assessment can be a promising approach, considering the complexity of human behaviour as well as the impressive amount of data involved. Neural networks, like the other methods mentioned above, have advantages and disadvantages. One advantage is represented by their flexibility, since they are an extension of the logistic regression; however, the flexibility has a price, as it poses more challenges to the interpretation of the model (Hamilton et al., 2014). Some authors have applied neural networks and conventional statistical techniques to predict criminal recidivism, and concluded that neural networks do not show any advantage (Caulkins et al., 1996). Others concluded that when the predictive performance of neural networks is compared to traditional methods, the results are inconsistent: two studies showed a better performance of neural networks, while two others did not find any significant difference (Liu et al., 2011).

5. Conclusion

We provided an overview of three classes of prediction models and cited some relevant application in the area of violence risk assessment. Different factors might play a role in the selection of one method over another, such as the purpose of the analysis, the homogeneity of the patient population, the type of predictors, the nature of the outcome variable, and the sample size.

Chapter VI

Risk of harm prediction at St. Joseph's Hospital

1. Introduction

In this chapter we study the risk of harm in inpatient psychiatric units at St Joseph's Hospital, with a specific focus on the prediction of the risk of harm. Towards this end we perform two important tasks: (1) We evaluate the predictability power of the Risk of Harm to Others Clinical Assessment Protocol (RHO CAP, Hirdes et al., 2011) for patients that have committed at least one violent during their stay in the hospital. To do so we augment the RHO CAP data with the hospital Safety Incident Reporting System (SIRS). (2) We determine the most relevant factors associated to harmful incidents and use them to develop an alternative predictive model. (3) We analyse the performance of the RHO CAP among patients who did not commit any aggression. In particular we find the distribution of time to first occurrence of violent events and show how it can be used to inform health care decision makers.

2. Data Collection

In our research we relied on two sources of data: the RHO CAP patients' records and the Safety Incident Reporting System (SIRS).

2.1 RHO CAP

The RHO CAP refers to an extended framework that assesses this risk in clinical settings, embracing various domains, without being constrained by the mode or the intention of the harm (Neufeld, Perlman & Hirdes, 2012). It is based on the Risk of Harm to Others (RHO) scale that uses a decision tree to classify harm based on specific behaviours and mental health symptoms, such as aggressive behaviours that occurred in the last three days, extreme behaviour, or violent acts (Neufeld, Perlman & Hirdes, 2012). The rating scale of the RHO follows the same logic used by other scales embedded in the RAI-MH, where higher scores are associated with more severe symptoms or impairments. The score ranges from 0 to 6, where 0 indicates no risk and 6 the highest risk of harm to others (Neufeld, Perlman & Hirdes, 2012). Based on the RHO scale, the RHO CAP generates a score that categorizes each patient in one of the f three categories: low risk, moderate risk, and high risk. The criteria adopted by the RHO CAP to categorize the residents are the following (Perlman et al., 2011; Neufeld, Perlman & Hirdes, 2012):

- *Low risk* = 0: includes all patients who receive a score between 0 and 2 on the RHO scale;

- *Moderate risk* = 1: includes all the patients how receive a score between 3 and 4 on the RHO scale;

- *High risk* = 2: includes all the patients who receive a score between 5 and 6 on the RHO Scale.

The RAI-MH is also equipped with guidelines that can be used by the staff when needed. Specifically, the main goal of the RHO CAP guidelines is to ensure the safety of all the individuals involved in different situations where the risk of harm to others is real, such as in a behavioural crisis or during an intervention to prevent or de-escalate an emergency (Neufeld, Perlman & Hirdes, 2012). For our research we used RHO CAP data in the years 2016 and 2017.

2.2 The Safety Incident Reporting System (SIRS)

The SIRS is a platform used to report all safety incidents that occur in psychiatric wards, covering the whole severity spectrum. For our research used SIRs records for the years 2016 and 2017. To report an incident a member of the clinical staff must fill in a standard form, immediately after the incident has occurred, detailing all the information related to the incident including patient's information, time and date of the incident, location, a brief description of the incident, and the level of severity attributed to the incident. Later on, the report is reviewed and the level of severity may be changes if necessary.

The first step in our data collection consisted in requesting a report from SIRS, showing all the incidents that happened at the St. Joseph's Hospital (Charlton and West 5th campuses) from January 2016 to December 2017. The report was pulled for 13 units and included schizophrenia, acute mental health, concurrent disorders, forensic psychiatry and addictions. The second step consisted in the anonymization of our dataset through the creation of a code assigned to each patient, to guarantee privacy and security. Based on the original SIRS report, we were able to retrospectively look at medical records, accessing the RAI-MH of patients, and extract the data needed.

The original report had 1045 incidents. The incident classification used to generate the report was "aggression to self/others". The types of incidents were organized in a column named "Specific Incident Type". We decided to remove two categories of incidents that were not needed for our research purposes. The:

- "abuse/threat/assault (physical) – victim" and
- "abuse/threat/assault (verbal) – victim"

These two categories have been excluded as we decided to focus on patients who committed aggressions.

The original report data on both the "Reported Incident Severity" and "Actual Incident Severity". We have considered the incident severity field as it is validated after the review of the SIRS record and as such is more accurate. The final sample of data obtained had 870 episodes of aggression incidents.

3. Data analysis

3.1 Age and Gender

We start by analysing the data for age and gender (see Table 3). The number of males who committed some type of aggression is almost double that of females. The mean (median) age of males is 46.37 (40.5), whereas the mean (median) for females is 41.63 (39).

	Mean	***Median***	***SD***	***N***
Age				
Male	46.37	40.5	19.58	224
Female	41.63	39	16.73	113
Age aggression incidents				
Male	50.34	52	19.89	565
Female	39.97	37	17.76	305

Table 3. Age and gender

The distribution of the incidents by gender confirms what has emerged in the literature review. Males are more prone to commit aggression when compared to females. In our sample, the number of males who committed violence is two times higher than the number of females, with 224 against 113. Males have committed 565 acts of aggression compared to 305 for females.

Males and females showed similar distributions of the RHO CAP scores (see Table 4). The Chi Squared Test revealed that the variables Gender and the RHO CAP are independent ($\chi 2 = 0.29$, df = 2; p-value = 0.865), leading us to conclude that there is no specific relationship between them.

Gender	**RHO CAP last RAI-MH before incident**				***p* value**
	0	**1**	**2**	**Total**	0.865
Males (total%) (% within males)	204 (23.45%) (36.11%)	130 (19.94%) (23.01%)	231 (26.55%) (40.88%)	565 (64.94%)	
Females (total%) (% within females)	114 (13.10%) (37.88%)	72 (8.27%) (23.61%)	119 (13.68%) (39.02%)	305 (35.06%)	
Total	318(36.55%)	202 (23.21%)	350 (40.2%)	870 (100%)	

```
            Large Sample
       Test Statistic     DF p-value
Chi Squared 0.29          2  0.865
```

Table 4. Relationship between gender and RHO CAP last assessment before incident

Figure 6 displays the count of the individual aggressions between males and females. The x-axis refers to the number of aggressions reported by each patient. As we can observe, the minimum of aggressions committed is 1, whereas the maximum refers to 28 episodes. The y-axis refers to the total counts for each category reported in the x-axis. The group of patients who committed only a single episode of aggression is the largest for both males and females (128 vs 60 episodes). It is interesting to note that as the number of episodes per patient increases, the difference between males and females related to the total counts of episodes decreases. We can conclude that while males commit more acts of violence the proportion of risk levels targeted doesn't seem to depend on the gender. This analysis provides more insights to the data showed in Table 4.

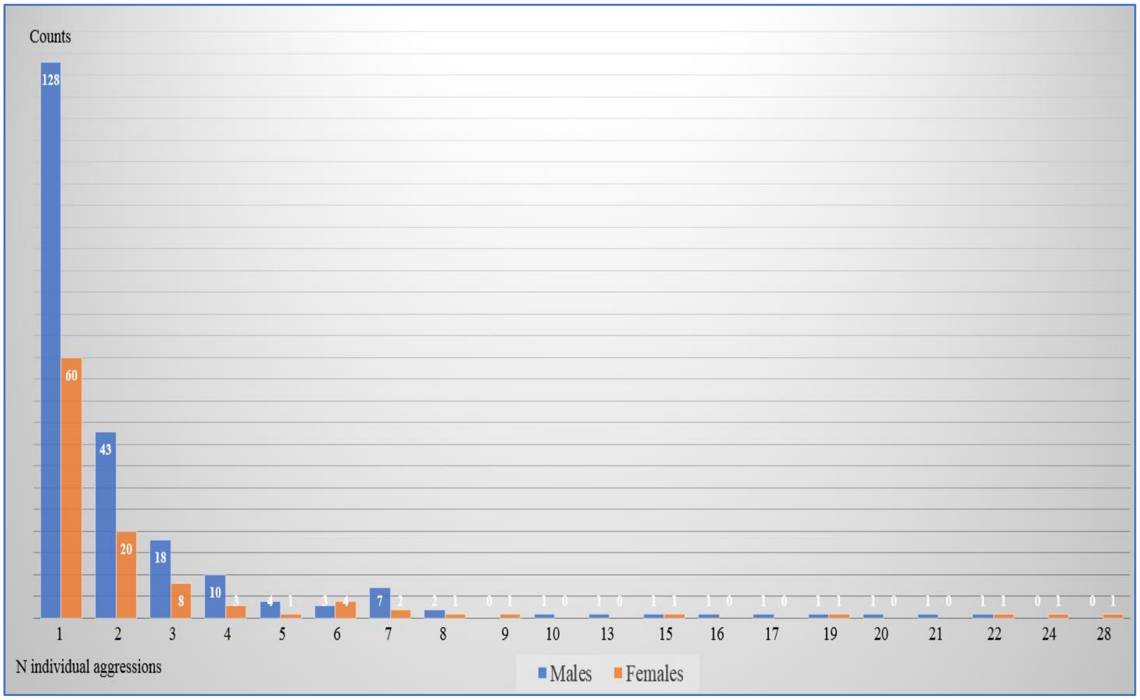


Figure 6. Distribution of individual aggressions between males and females

In order to have more insights about how differently males and females are involved with aggression, we used:

1 – a contingency table to describe the interrelation between the severity of the incidents reported (Incident Severity for Reporting) and the gender (Table 5);

2 – a correlational analysis to describe the relationship between the severity of the incidents reported (Incident Severity for Reporting) and the age (Figure 9).

3.1.1 Gender and Incident Severity for Reporting

A significant relationship between these two variables has emerged ($\chi 2 = 14.385$, df = 4; p-value = 0.006). The highest number of incidents belongs to Level 2 (No Harm), with males responsible for 417 (47.93%) episodes while females are responsible for 192 (62.95%) (Table 5 and Figure 7). The distribution of the other levels of severity between males and females is similar (Table 5 and Figure 8), with a small exception: aggressions that caused mild harm (Level 3) were more frequent in females (28.2% vs 19.82%).

Incident Severity for Reporting	**Gender**		**Total**	***p-value***
	Males	**Females**		*0.006*
Severity Level 1-Near Miss (total %) (% within gender category)	24 (2.76%) (4.25%)	15 (1.72%) (4.92%)		39
Severity Level 2-No Harm Incident (total %) (% within gender category)	417 (47.93%) (73.81%)	192 (22.07%) (62.95%)		609
Severity Level 3-Harmful Incident - Mild Harm (total %) (% within gender category)	112 (12.87%) (19.82%)	86 (9.89%) (28.20%)		198
Severity Level 4-Harmful Incident - Moderate Harm (total %) (% within gender category)	12 (1.38%) (2.12%)	10 (1.15%) (3.28%)		22
Severity Level 5-Harmful Critical Incident – Severe Harm (total %) (% within gender category)	0	0		0
Severity Level 6-Harmful Critical Incident – Death (total %) (% within gender category)	0	*X**		*X**
Total	565	305	870	

```
              Large Sample
           Test Statistic    DF p-value
Chi Squared 14.385           4   0.006
```

**For privacy concerns, the data is omitted for samples less than or equal to 3*

Table 5. Relationship between severity of incidents and gender

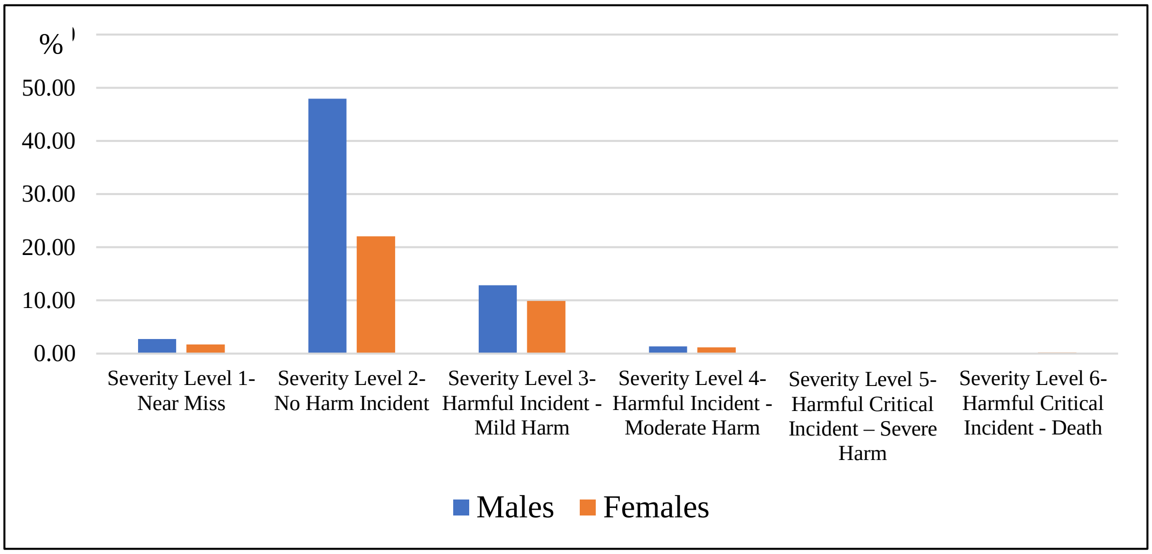


Figure 7. Incident Severity for Reporting - Last RAI-MH before aggression - Total %

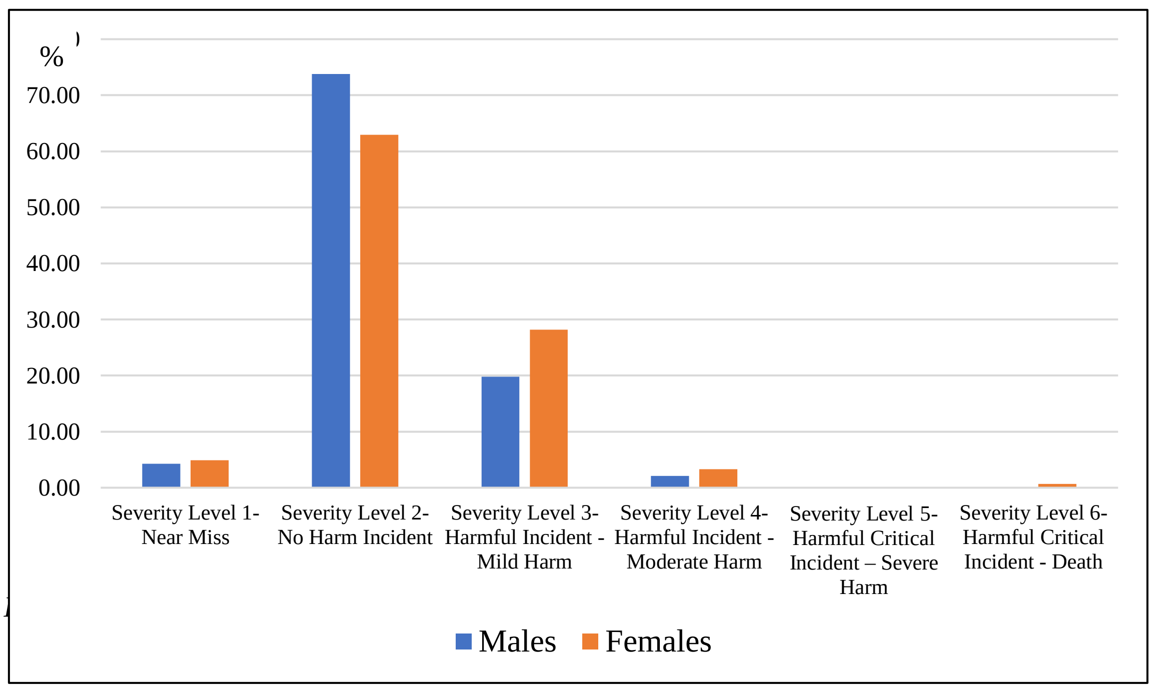
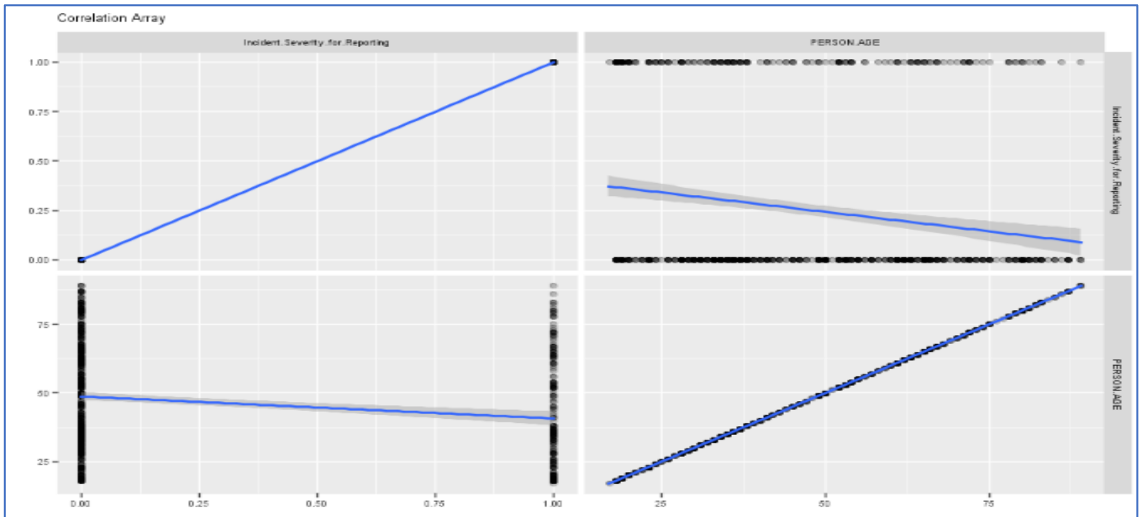


Figure 8. Incident Severity for Reporting - Last RAI-MH before aggression - % within gender category

3.1.2 Age and Incident Severity for Reporting: correlation analysis




```
                                                  Spearman's rank correlation
rho

                                PERSON.AGE                  Incident.Severity.for.Reporting
          PERSON.AGE       cor  1                           -0.1911
                   N            870                         870
              stat**                                        130720343
             p-value                                        0.0000
---------------------------
Incident.Severity.for.Reporting  cor -0.1911                 1
                   N            870                         870
              stat**            130720343
             p-value            0.0000
---------------------------
     ** S
     HA: two.sided
```

Figure 9. Correlation analysis between age and incident severity

A correlation analysis between Incident Severity for Reporting and Age was performed (Figure 9). Given the nature of the two variables, we selected a non-parametrical method. The Spearman correlation is a non-parametric measure of rank correlation between two variables, which assesses the statistical dependence between the ranking of two variables. In contrast with the Pearson's coefficient, it assesses monotonic relationships, testing for linear relationships. Its values range from -1 to 1. The hypotheses have been set as follows:

- H0: There is no correlation between the variables
- Ha: There is a correlation between the variables

The interval was fixed at 95%, with a p-value lower than 0.05, confirming a significant relationship between the variables. The coefficient is -0.1911, describing a very weak negative relationship. A lower age is correlated to higher levels of severity of the incidents. This result finds parallels with the literature, as it confirms that younger patients are more prone to commit more severe aggressions.

3.2 Incidents Severity for Reporting and RHO CAP

We were interested in studying the relationship between the RHO CAP and the severity of the incidents reported. This analysis gives an overview of the performance of the RHO CAP in predicting the risk of harm, as we compared the algorithm score with the actual severity of the incident reported by the clinical staff. This is in line with the main intent of our study, which is to investigate the performance of this algorithm in predicting violent behaviours of residents in the inpatient setting. The contingency results (Table 6) show that the highest number of incidents are categorized under Level 2 (No Harm Incident) and Level 3 (Harmful Incident - Mild Harm), and reveal a significant association between the RHO CAP and levels of reported severity ($\chi2 = 22.046$, df = 8; p-value = 0.005). In evaluating the performance of the RHO CAP, we assumed that

- a RHO CAP = 0 corresponds to Severity Levels 1 and 2
- a RHO CAP = 1 corresponds to Severity Levels 3 and 4

- a RHO CAP = 2 corresponds to Severity Levels 5 and 6

From the data showed in Table 4 we can assume that, during the time considered, the RHO CAP has predicted with more efficiency cases of aggression with a low or mild level of severity, corresponding to Severity Levels 1 and 2. The RHO CAP predicted correctly the risk of harm of the 31% of cases reported, with an accuracy for the cases not at risk of harm of 24% against a 7% of accuracy for the cases at risk of harm (mild or severe risk).

Incident Severity for Reporting	**RHO CAP last RAI-MH before incident**				***p*-value**
	0	1	2	Total	0.005
Severity Level 1 – Near Miss (total %) (% within level of severity)	17 (1.95%) (43.59%)	10 (1.15%) (25.64%)	12 (1.38%) (30.77%)	39 (4.48%)	
Severity Level 2 – No Harm (total %) (% within level of severity)	200 (22.99%) (32.84%)	136 (15.63%) (22.33%)	273 (31.38%) (44.83%)	609 (70%)	
Severity Level 3 – Mild Harm (total %) (% within level of severity)	87 (10%) (43.94%)	51 (5.86%) (25.76%)	60 (6.7%) (30.30%)	198 (22.76%)	
Severity Level 4 – Moderate Harm (total %) (% within level of severity)	13 (1.49%) (59.09%)	4 (0.46%) (18.18%)	5 (0.57%) (22.73%)	22 (2.53%)	
Severity Level 5 – Severe Harm (total %) (% within level of severity)	0	0	0	0	
Severity Level 6 – Death	*X**	*X**	0	*X**	

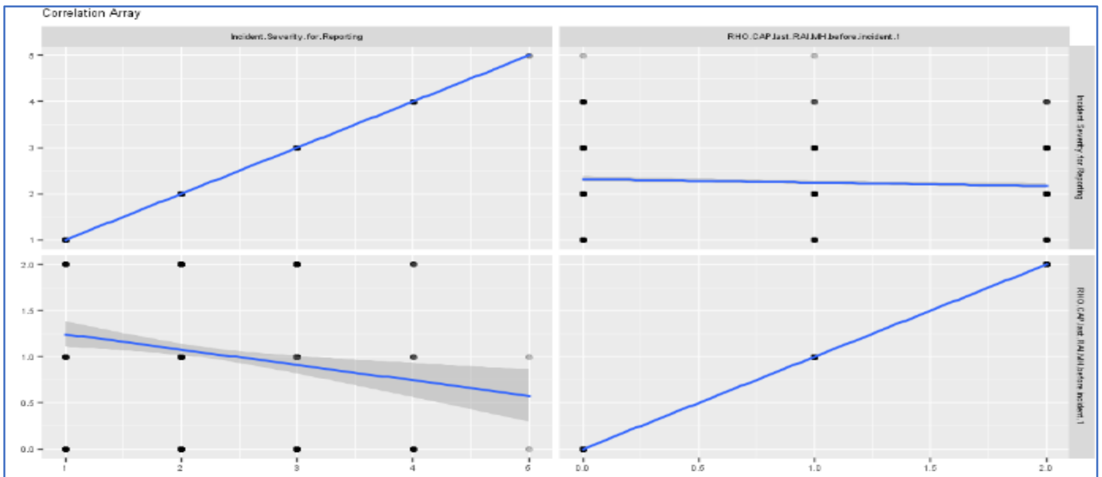
```
          Large Sample
     Test Statistic     DF p-value
Chi Squared 22.046      8   0.005
```

**For privacy the data is omitted for samples less than or equal to 3*

Table 6. Relationship between severity of incidents and RHO CAP last assessment

According to Table 6, the RHO CAP shows an overall accuracy in predicting the risk of harm of approximately 31.26%: this figure is obtained by summing up the percentages of the quadrants that indicate the correspondence between the RHO CAP score and its relative Level of Severity. The accuracy for the cases not at risk of harm is 24.94%, whereas the accuracy for the cases at risk of harm (mild or severe risk) is 6.32%.

To investigate the relationship between the RHO CAP (generated by the last RAI-MH before the aggression incident) and the Incident Severity reported by the staff, a correlation analysis was performed through a non-parametrical method (Figure 10).




```
                                          Spearman's rank correlation rho

                                     RHO.CAP.last.RAI.MH.before.incident.1 Incident.Severity.for.Reporting
RHO.CAP.last.RAI.MH.before.incident.1     cor 1                                     -0.1098
                                  N           870                                   870
                              stat**                                                121802103
                             p-value                                                0.0012
---------------------------
      Incident.Severity.for.Reporting     cor -0.1098                               1
                                  N           870                                   870
                              stat**          121802103
                             p-value          0.0012
---------------------------
      ** S
      HA: two.sided
```

Figure 10. Correlation analysis between RHO CAP and incident severity

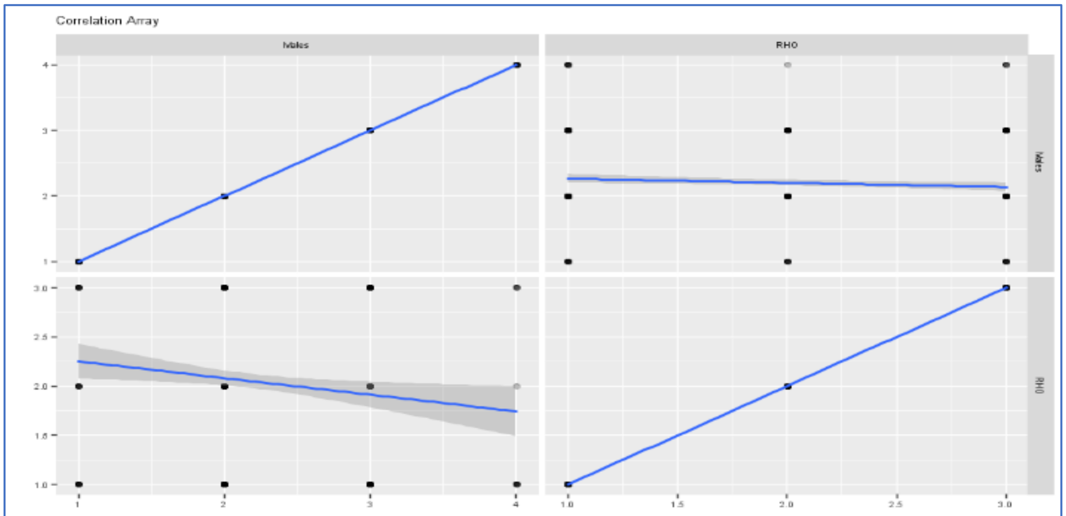
Our hypotheses are:

- H0: There is no correlation between the variables
- Ha: There is a correlation between the variables.

The interval was fixed at 95%, with a p-value lower than 0.05, confirming a significant relationship between the variables. The coefficient is -0.1098, describing a very weak negative relationship. Higher scores of the RHO CAP tend to be correlated with lower levels of severity of the incidents. This result is interesting, as it shows that higher scores of the RHO CAP generated by the last assessment (RAI-MH) are not necessarily associated with more severe incidents. We can say that the RHO CAP generates scores that can be interpreted as false positives: it predicts a high risk of harm when the severity of the incident is very low.

A further correlation analysis for males (Figure 11) and females (Figure 12) was conducted, revealing similar results:

Males

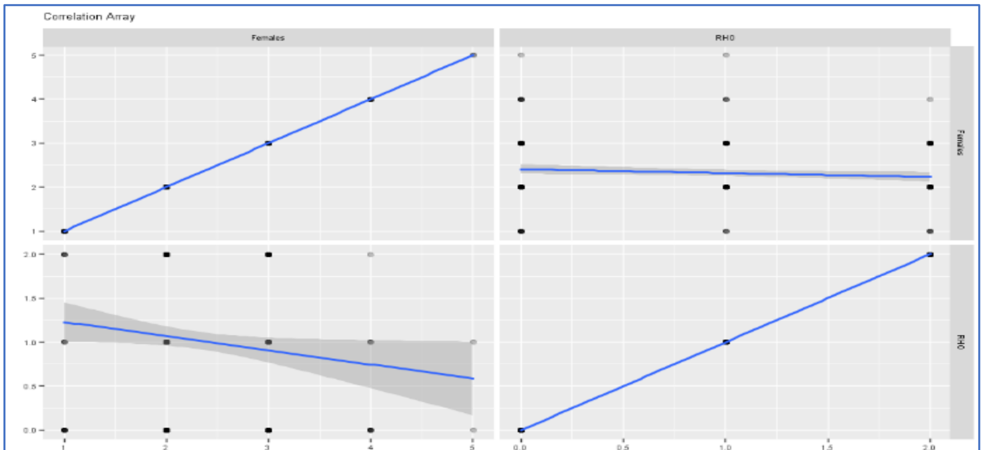



```
                                                        Spearman's rank correlation
rho

            Males    RHO
Males    cor 1        -0.1064
           N 565      565
      stat**          33258579
     p-value          0.0114
------------
  RHO    cor -0.1064  1
           N 565      565
      stat** 33258579
     p-value 0.0114
------------
      ** S
      HA: two.sided
```

Figure 11. Correlation analysis between males and RHO CAP

Females




```
                                          Spearman's rank correlation
rho

               RHO     Females
    RHO    cor 1       -0.1158
             N 305     305
        stat**         5276191
       p-value         0.0433
--------------
Females    cor -0.1158 1
             N 305     305
        stat** 5276191
       p-value 0.0433
--------------
      ** S
      HA: two.sided
```

Figure 12. Correlation analysis between females and RHO CAP

In both cases the p-value is lower than 0.05. We found very weak negative correlation (-0.1064 for males, -0.1158 for females) between the RHO CAP and the levels of severity of the incidents, which is consistent with the previous result.

3.3 Education and RHO CAP

Table 7 displays the relationship between levels of education and the RHO CAP scores, with a significant association between these two variables ($\chi 2 = 49.241$, df = 14; p-value < 0.001). Residents with lower levels of education were responsible for most of the reported aggressive incidents (Table 7). In fact, the total percentage of episodes committed by those who have an educational background that goes from “no education” to “high school” is 66.31%. We can assume that education might play a role in mitigating aggressive behaviours within the inpatient settings. In addition, residents with lower levels of education are identified as patients at higher risk of harm to others by the algorithm (RHO CAP = 2). Individuals with no education at all (no schooling) are responsible for the highest number of incidents (27.93%). The highest RHO CAP score (2 = High risk of harm to others) has been reported by individuals with no education (no schooling). However, we have a significant percentage of missing data, and it was not possible to identify the level of education for the 21.38% of the residents.

Level of Education	**RHO CAP last RAI-MH before incident**				***p* value**
	0	1	2	Total	< 0.001
No schooling (total %) (% within level of education)	89 (9.89%) (36.63%)	62 (7.13%) (25.51%)	95 (10.92%) (39.09%)	243 (27.93%)	
8th grade/less (total %) (% within level of education)	21 (2.41%) (36.84%)	7 (0.8%) (12.28%)	29 (3.33%) (50.88%)	57 (6.55%)	
9-11 grades (total %) (% within level of education)	50 (5.75%) (46.30%)	29 (3.33%) (26.85%)	29 (3.33%) 26.85%)	108 (12.41%)	
High School (total %) (% within level of education)	76 (8.74%) (44.97%)	40 (4.6%) (23.67%)	53 (6.09%) (31.36%)	169 (19.42%)	
Technical or trade school (total %) (% within level of education)	4 (0.46%) (23.53%)	3 (0.34%) (17.65%)	10 (1.15%) (58.82%)	17 (1.95%)	
Diploma/Bachelor’s Degree (total %) (% within level of education)	12 (1.38%) (54.55%)	3 (0.34%) (13.64%)	7 (0.8%) (31.82%)	22 (2.53%)	
Some Collage/University (total %) (% within level of education)	24 (2.76%) (30.77%)	26 (2.99%) (33.33%)	28 (3.22%) (35.90%)	78 (8.96%)	
Unknown (total %) (% within level of education)	45 (5.17%) (25.57%)	32 (3.68%) (18.18%)	99 (11.38%) (56.25%)	176 (21.38%)	

```
             Large Sample
        Test Statistic     DF p-value
Chi Squared 49.241         14 <0.001
```

Table 7. relationship between level of education and RHO CAP

3.4 Education and Incident Severity for Reporting

The level of education also appeared related to the severity of the aggression episodes ($\chi 2$ = 66.198, df = 28; p-value < 0.001). Individuals with less education seem more at risk of committing more severe episodes of aggression, compared to more educated ones (Table 8).

	Incident Severity for Reporting						p-value
	Severity Level 1-Near Miss	Severity Level 2-No Harm Incident	Severity Level 3-Harmful Incident - Mild Harm	Severity Level 4-Harmful Incident - Moderate Harm	Severity Level 5-Harmful Critical Incident – Severe Harm	Severity Level 6- Harmful Critical Incident - Death	< 0.001
Level of Education							
No schooling (total %) (% within level of education)	7 (0.80%) (2.88%)	174 (20%) (71.60%)	53 (6.09%) (21.81%)	9 (1.03%) (3.70%)	0	0	243
8^{th} grade/less	0	43 (4.94%) (75.44%)	14 (1.61%) (24.56%)	0	0	0	57
9-11 grades	8 (0.92%) (7.41%)	65 (7.47%) (60.19%)	32 (3.68%) (29.63%)	3 (0.34%) (2.78%)	0	0	108
High School	12 (1.38%) (7.10%)	102 (11.72%) (70.36%)	46 (5.29%) (27.22%)	8 (0.92%) (4.73%)	0	1 (0.11%) (0.59%)	169
Technical or trade school	1 (0.11%) (5.88 %)	15 (1.72%) (88.24%)	1 (0.11%) (5.88%)	0	0	0	17
Diploma/Bachelor's Degree	2 (0.23%) (9.09%)	13 (1.49%) (59.09%)	6 (0.69%) (27.27%)	0	0	1 (0.11%) (4.55%)	22
Some Collage/University	6 (0.69%) (7.69%)	51 (5.86%) (65.38%)	20 (2.30%) (25.64%)	1 (0.11%) (1.18%)	0	0	78
Unknown	3 (0.34%) (1.70%)	146 (16.78%) (65.38%)	26 (2.99%) (14.77%)	1 (0.11%) (0.57%)	0	0	176
Total	39	609	198	22	0	2	870

```
             Large Sample
        Test Statistic    DF p-value
Chi Squared 66.198        28 <0.001
```

8. Relationship between education and incident severity

3.5 Diagnosis and RHO CAP

One of the most important predictors of violence in the literature is the patient's diagnosis. Among the most frequent disorders associated with inpatient aggression we can mention psychotic disorders, personality disorders and substance abuse disorder. Other relevant factors include history of violence, exposition to violence and early onset of criminal behaviour. We found a significant relationship between the last diagnosis received before the aggression episode and the risk of harm ($\chi 2 = 208.455$, $df = 60$; $p\text{-value} < 0.001$). The most common (provisional) diagnoses received in our sample belongs to the following categories:

- Schizophrenia Spectrum and Other Psychotic Disorders (33.68%)
- Neurocognitive Disorders (19.08%)
- Personality Disorders (9.54%)
- Bipolar and Related Disorders (5.75%)

Within each diagnostic category, the highest risk of harm (RHO CAP = 2) was predicted for patients who received a diagnosis of Trauma and Stressor-Related Disorders, Obsessive-Compulsive and Related Disorders, Neurocognitive Disorders, and Schizophrenia Spectrum and Other Psychotic Disorders (Table 9, Figure 13). However, in our sample 19.8% of patients have not received a diagnosis during the last RAI-MH completion. The missing data might have an impact on final conclusions. Our results confirm similar findings in the literature. Neurocognitive disorders represent the second most frequent diagnostic category in our sample and the category with the highest percentage of RHO CAP scores equal to 2 (63.25%), if we exclude Trauma and Stressors-

Related Disorders and Obsessive-Compulsive and Related Disorders, due to their small number of diagnoses. In the literature, the relationship between problems in neurocognitive domain and inpatient violence has been frequently studied in patients affected by schizophrenia (Reinharth et al., 2014; Bulgari et al., 2017) or schizoaffective disorder (O'Reilly et al., 2015). In our sample the diagnosis related to neurocognitive disorder was not secondary to other disorders. More investigations are needed to further explore the relationship of this diagnostic category and inpatient violence.

Most important diagnosis before incident (DSM-V)	**RHO CAP last RAI-MH before**				***p* value**
	0	**1**	**2**	**Total**	< 0.001
Bipolar and Related Disorders	20 (2.3%) (40%)	22 (2.53%) (44%)	8 (0.92%) (16%)	50 (5.75%)	
Anxiety Disorders	1 (0.11%) (100%)	0	0	1 (0.11%)	
Depressive Disorders	7 (0.8%) (58.33%)	4 (0.46%) (33.33%)	1 (0.11%) (8.33%)	12 (1.38%)	
Disruptive, Impulse-Control, and Conduct Disorders	5 (0.57%) (83.33%)	0	1 (0.11%) (16.67%)	6 (0.69%)	
Medication-Induced Movement Disorders and Other Adverse Effects of Medication	5 (0.57%) (83.33%)	1 (0.11%) (16.67%)	0	6 (0.69%)	
Neurocognitive Disorders	30 (3.45%) (18.07%)	31 (3.56%) (18.67%)	105 (17.24%) (63.25%)	166 (19.08%)	
Neurodevelopmental Disorders	22 (2.53%) (52.38%)	8 (0.92%) (19.05%)	12 (1.38%) (28.57%)	42 (4.83%)	
Obsessive-Compulsive and Related Disorders	1 (0.11%) (20%)	0	4 (0.46%) (80%)	5 (0.57%)	
Personality Disorders	41 (4.71%) (49.4%)	22 (2.53%) (26.51%)	20 (2.3%) (24.10%)	83 (9.54%)	
Schizophrenia Spectrum and Other Psychotic Disorders	106 (12.18%) (36.18%)	76 (8.73%) (25.94%)	111 (12.76%) (37.88%)	293 (33.68%)	
Substance-Related and Addictive Disorders	15 (1.72%) (62.50%)	4 (0.46%) (16.67%)	5 (0.57%) (20.83%)	24 (2.76%)	
Trauma- and Stressor-Related Disorders	0	0	9 (1.03%) (100%)	9 (1.03%)	
Paraphilic Disorders	1 (0.11%) (100%)	0	0	1 (0.11%)	
Missing diagnosis	64 (7.36%) (37.21%)	34 (3.9%) (19.77%)	74 (8.5%) (43.02%)	172 (19.8%)	

```
             Large Sample
        Test Statistic    DF p-value
Chi Squared 208.455       60 <0.001
```

Table 9. Relationship between diagnosis and RHO CAP

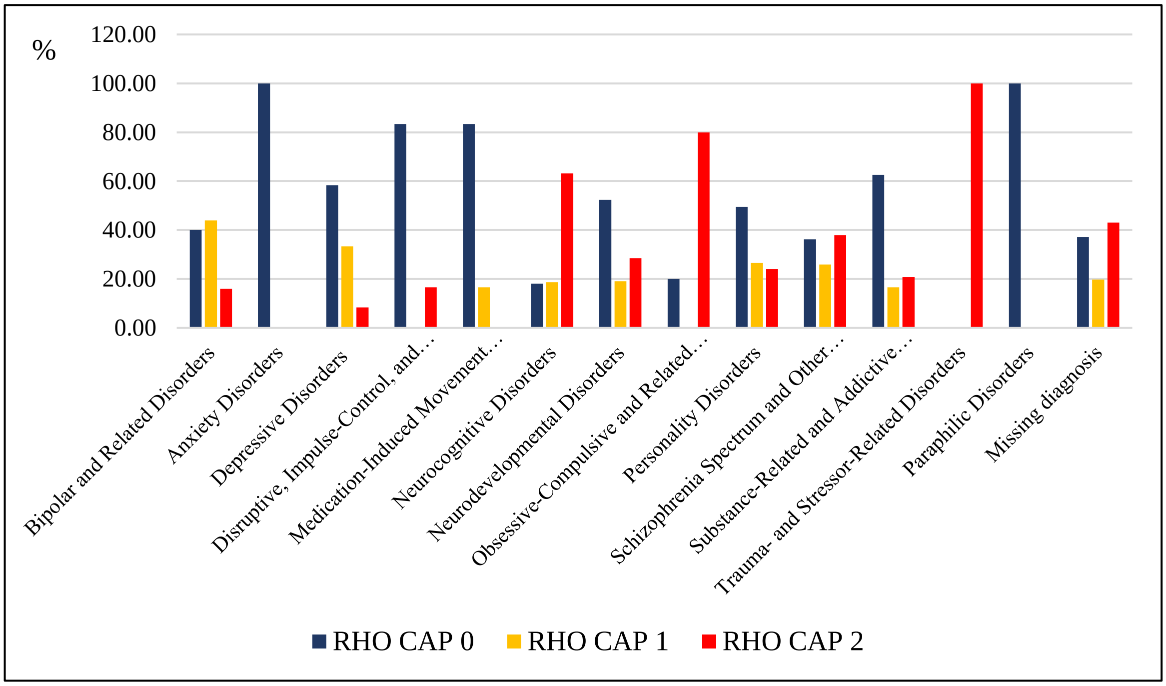


Figure 13. Most important diagnosis before aggression – % within each diagnostic category

3.6 RHO CAP and capacity of insight into mental health problem

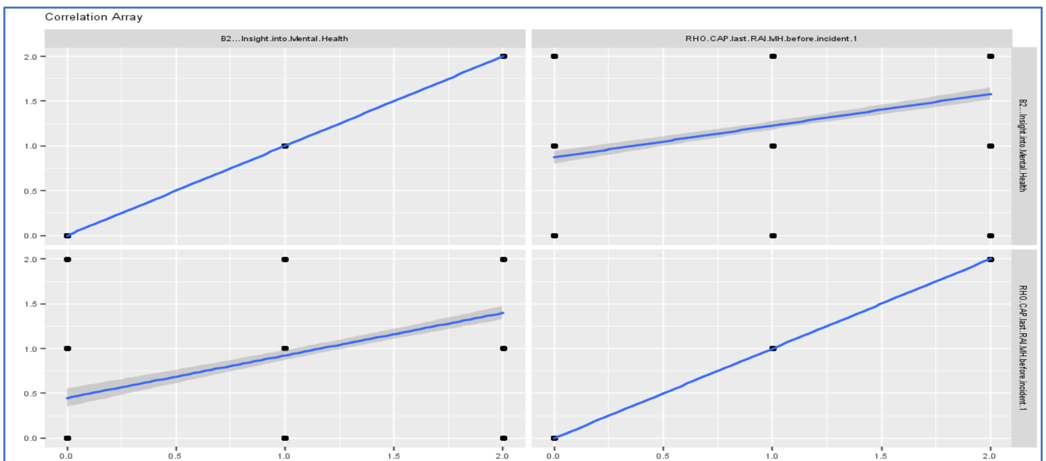
In this context, the capacity of insight to mental health problems refers to the general capacity of a patient to be aware and conscious of a problem. This construct is identified by three levels of insight: full (0), limited (1) and absent (2). Our data reveal a significant relationship between patients' deteriorated levels of insight into mental health and higher risk of harm to others ($\chi2 = 238.503$, $df = 4$; $p\text{-value} < 0.001$). In other words, patients with limited ability to understand their problem are identified as more at risk of perpetrate aggression against others. Table 10 shows that the highest percentage of patients who received a RHO CAP score of 2 have been assessed with the lowest degree of mental health insight (None).

	RHO CAP last assessment before incident				*p-value*
Insight into mental health problem	**0**	**1**	**2**	**Total**	< 0.001
Full	82 (9.43%) (25.79%)	63 (7.24%) (19.81%)	25 (2.87%) (7.86%)	170 (19.54%)	
Limited	161 (18.51%) (79.70%)	98 (11.26%) (48.51%)	63 (7.24%) (31.19%)	322 (37.01%)	
None	75 (8.62%) (24.59%)	41 (4.71%) (13.44%)	262 (30.11%) (85.90%)	378 (43.45%)	

```
               Large Sample
            Test Statistic     DF p-value
Chi Squared 238.503            4  <0.001
```

Table 10. Relationship between RHO CAP and Insight into Mental Health

A correlation analysis between the RHO CAP and Insight into Mental Health has also been performed and reported in Figure 14.




```
                                   Spearman's rank correlation rho

                                        RHO.CAP.last.RAI.MH.before.incident.1 B2...Insight.into.Mental.Health
RHO.CAP.last.RAI.MH.before.incident.1    cor 1                                    0.4328
                                    N        870                                      870
                               stat**                                             62248450
                              p-value                                             0.0000
--------------------------
      B2...Insight.into.Mental.Health    cor 0.4328                               1
                                    N        870                                      870
                               stat**        62248450
                              p-value        0.0000
--------------------------
     ** S
     HA: two.sided
```

Figure 14. Correlation analysis between RHO CAP and Insight into Mental Heath

Hypotheses and confidence intervals have been set consistently with the criteria descripted in our previous correlation analyses. The result of this analysis shows a moderate positive correlation between the RHO CAP and the insight to mental health. This is in line with our expectations, because it shows that higher values of the RHO CAP are correlated with lower levels of mental health insight.

3.7 Aggressive Behaviour Scale (ABS)

We investigated the Aggressive Behaviour Scale, and its relationship with other two variables of interests: the RHO CAP and the Levels of Severity of incidents. The purpose of the ABS is to provide a quick picture, easily comprehensible, of the severity of the aggressive behaviour of an individual. It finds applications for different purposes, including quality monitoring and research (Perlman & Hirdes, 2008). As a validated measure of clinical relevance, we decided to investigate its relationship with three other clinically relevant variables in our study: Gender, the RHO CAP, and “Diagnosis.”.

3.7.1 ABS and Gender

Table 11 and Figure 15 provide an overview of how the ABS scores are distributed among males and females ($\chi 2 = 7.957$, df = 3; p-value = 0.047). The distribution of the ABS scores with each gender category is similar. According to this scale, females have demonstrated to be slightly more prone to show very severe aggressive behaviours than males (36.07% vs 29.73%). Males have shown a higher percentage of moderately severe ABS scores (17.17% vs 10.82%).

	MALES		FEMALES		*p-value*
	Total %	% within males	Total %	% within females	0.047
None (0)	19.20%	29.56%	10.69%	30.49%	
Moderate (1-2)	11.15%	17.17%	3.79%	10.82%	
Severe (3-5)	15.29%	23.54%	7.93%	22.62%	
Very severe (6+)	19.31%	29.73%	12.64%	36.07%	
Total	64.94%	100.00%	35.06%	100.00%	

```
             Large Sample
        Test Statistic     DF p-valu
Chi Squared 7.957          3  0.047
```

Table 11. Distribution of ABS scores among males and females

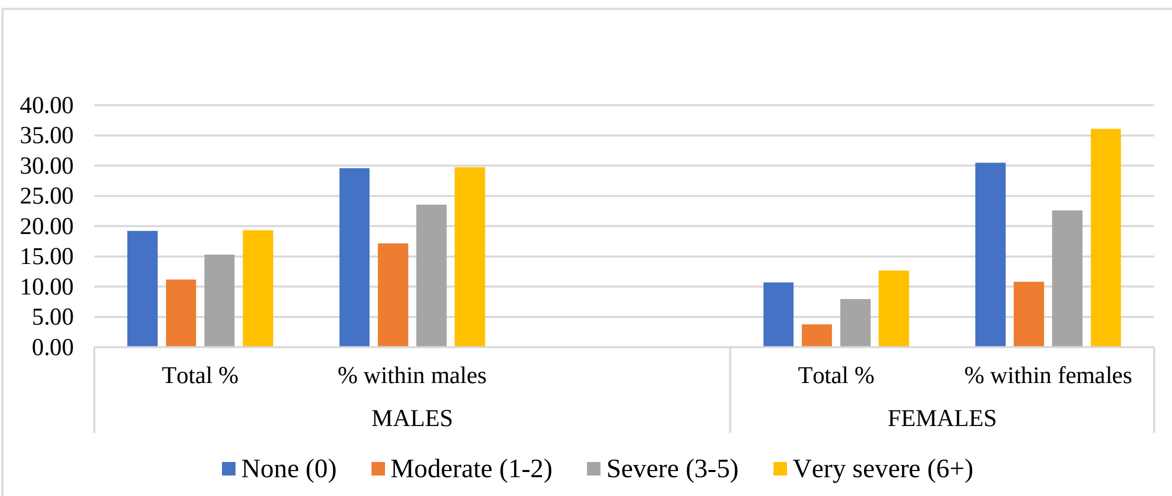


Figure 15. ABS by Gender

3.7.2 ABS and RHO CAP

Analysing the interaction between the ABS and the RHO CAP (Table 12, Figure 16), we obtained a significant relationship between the variables ($\chi 2$ = 225.124, df = 6; p-value <

0.001). About 66.92% of the patients who received an ABS score of 0 (None), have been identified as not at risk of harm to others (RHO CAP = 0). On the opposite pole, 46.04% of patients who received a sever ABS score and 66.91% of patients who received a very severe ABS score, have also been identified at high risk of harm to others (RHO CAP = 2).

	RHO CAP				
	0	**1**	**2**	**Total**	***p-value*** < 0.001
ABS Score					
None [0] (total %) (% ABS level)	174 (20%) (66.92%)	53 (6.09%) (22.38%)	33 (3.79%) (12.69%)	260 (29.89%)	
Moderate [1-2] (total %) (% ABS level)	56 (6.44%) (43.08%)	36 (4.14%) (27.69%)	38 (4.37%) (29.23%)	130 (14.94%)	
Severe [3-5] (total %) (% ABS level)	51 (5.86%) (25.25%)	58 (6.67%) (28.71%)	93 (10.69%) (46.04%)	202 (23.22%)	
Very severe [6+] (total %) (% ABS level)	37 (4.25%) (13.31%)	55 (6.32%) (19.78%)	186 (21.38%) (66.91%)	278 (31.95%)	
Total	318 (36.55%)	202 (23.22%)	350 (40.23%)	870	

	Large Sample Test Statistic	DF	p-value
Chi Squared	225.124	6	<0.001

Table 12. Relationship between RHO CAP and ABS

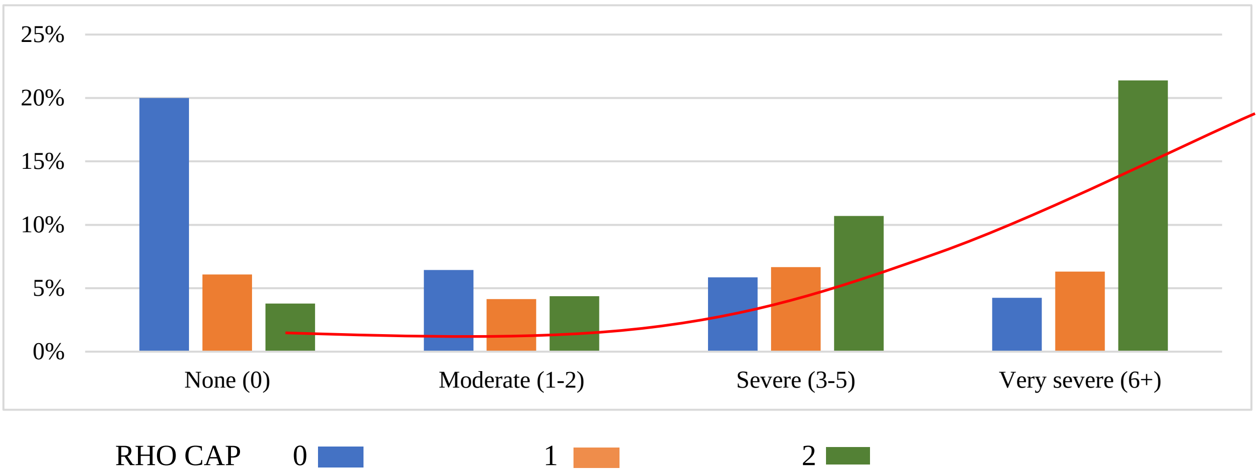
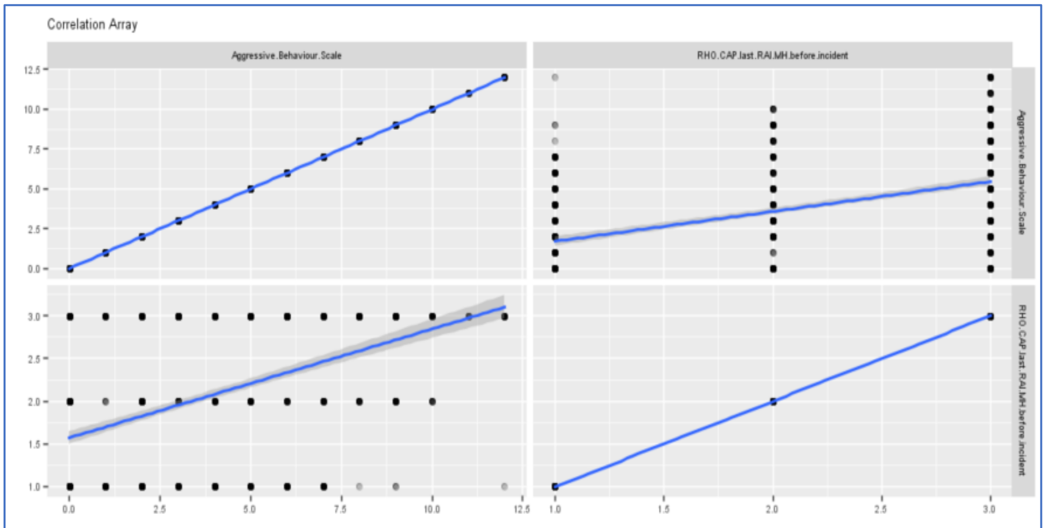


Figure 16. ABS and RHO CAP

The correlation test between the ABS and the RHO CAP showed a significant result (p-value <0.05), in line with the previous analysis (Figure 17).




```
                                            Spearman's rank correlation rho

                                   RHO.CAP.last.RAI.MH.before.incident Aggressive.Behaviour.Scale
RHO.CAP.last.RAI.MH.before.incident    cor 1                                 0.5004
                          N                870                               870
                     stat**                                                  54829824
                    p-value                                                  0.0000
--------------------------
         Aggressive.Behaviour.Scale    cor 0.5004                            1
                          N                870                               870
                     stat**                54829824
                    p-value                0.0000
--------------------------
    ** S
    HA: two.sided
```

Figure 17. Correlation analysis between RHO CAP and ABS

The Spearman coefficient is 0.5004, showing a moderate positive correlation. Higher values of risk of harm, detected by the algorithm, tend to be associated with higher levels of aggressive behaviours according to the ABS. Similarly, lower values of the RHO CAP tend to be associated with lower or minimal levels of aggressive behaviours according to the ABS.

3.7.3 ABS and Diagnosis

We found a significant relationship between the ABS and the last most important diagnosis received before an aggression incident (p-value <0.001). The diagnostic categories showing the highest ABS scores are "Schizophrenia Spectrum and Other Psychotic Disorders" and "Neurocognitive Disorders". The percentage of missing data in this case is significantly high, peaking to 19.77% (Table 13).

Most important diagnosis before incident (DSM-V)			**Aggressive Behaviour Scale**		***(p-value) <0.001)***
	None	**Moderate**	**Severe**	**Very Severe**	**Total**
Anxiety Disorders (total %) (%within diagnostic category)	1 (0.11%) (100 %)	0	0	0	1 (0.11%) (100 %)
Bipolar and Related Disorders (total %) (%within diagnostic category)	18 (2.07%) (36%)	6 (0.69%) (12%)	13 (1.49%) (26%)	13 (1.49%) (26%)	50 (5.75%) (100%)
Depressive Disorders (total %) (%within diagnostic category)	7 (0.80%) (58.33%)	2 (0.23%) 16.67%)	1 (0.11%) (8.33%)	2 (0.23%) (16.67%)	12 (1.38%) (100%)
Disruptive, Impulse-Control, and Conduct Disorders (total %) (%within diagnostic category)	0	0	6 (0.69%) (100%)	0	6 (0.69%) (100%)
Neurocognitive Disorders (total %) (%within diagnostic category)	27 (3.10%) (16.27%)	30 (3.45%) (18.07%)	40 (4.60%) (24.10%)	69 (7.93%) (41.57%)	166 (19.08%)
Neurodevelopmental Disorders (total %) (%within diagnostic category)	3 (0.34%) (7.14%)	4 (0.46%) (9.52%)	12 (1.38%) (28.57%)	23 (2.64%) (54.17%)	42 (4.83%) (100%)
Obsessive-Compulsive and Related Disorders (total %) (%within diagnostic category)	1 (0.11%) (20%)	0	0	4 (0.46%) (80%)	5 (0.57%) (100%)
Personality Disorders (total %) (%within diagnostic category)	37 (4.25%) (44.58%)	14 (1.61%) (16.87%)	18 (2.07%) (21.69%)	14 (1.61%) (16.87%)	83 (9.54%) (100%)
Schizophrenia Spectrum and Other Psychotic Disorders (total %) (%within diagnostic category)	109 (12.53%) (378.20%)	38 (4.37%) 12.97%)	71 (8.16%) (24.23%)	75 (8.62%) (25.60%)	293 (33.68%) (100%)
Substance-Related and Addictive Disorders (total %) (%within diagnostic category)	13 (1.49%) (54.17%)	4 (0.46%) (16.67%)	3 (0.34%) (12.50%)	4 (0.46%) (16.67%)	24 (2.76%) (100%)
Trauma- and Stressor-Related Disorders (total %) (%within diagnostic category)	0	0	0	9 (1.03%) (100%)	9 (1.03%) (100%)
Medication-Induced Movement Disorders and Other Adverse Effects of Medication (total %) (%within diagnostic category)	0	6 (0.69%) (100%)	0	0	6 (0.69%) (100%)
Paraphilic Disorders (total %) (%within diagnostic category)	1 (0.11%) (100%)	0	0	0	1 (0.11%) (100%)
Missing data	43 (4.94%) (25%)	26 (2.99%) 15.12%)	38 (4.37%) (22.09%)	65 (7.47%) (37.79%)	172 (19.77%) (100%)
Total	260 (29.89%)	130 (14.94%)	202 (23.22%)	278 (31.95%)	870 (100%)

Table 13. Relationship between ABS and Diagnosis

The bar graphs in Figures 8 and 9 display the content reported in the contingency table, showing the percentages of the ABS scores for each diagnostic category. Figure 18 shows how the ABS scores are distributed within the total. If we consider the total of percentages, the diagnostic categories that are associated to the highest scores of the ABS are the psychotic disorders (and the neurocognitive disorders. This data supports the previous analysis addressing the relationship between the RHO CAP and the diagnosis (Table 9). A RHO CAP of 2 was triggered by the 17.24% of aggressions associated with a diagnosis of neurocognitive disorders and by the 12.76% of aggressions associated with a diagnosis of psychotic disorder.

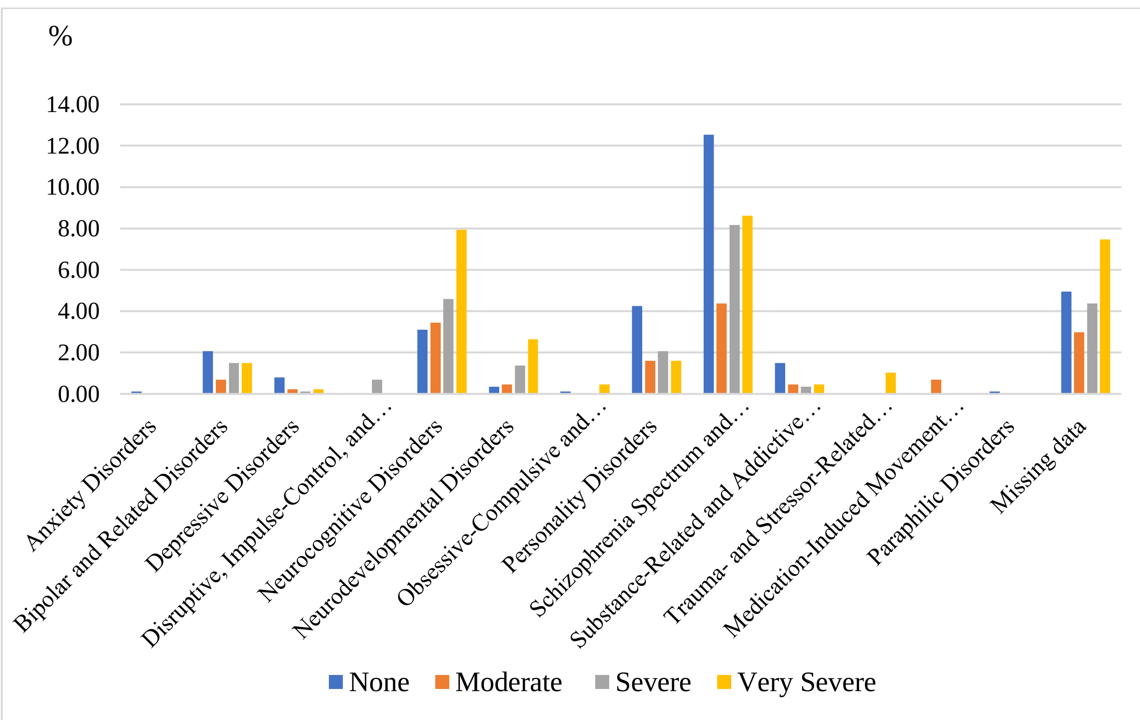


Figure 18. Aggressive Behaviour Scale by last diagnosis before aggression incident – Total %

3.8 Spatial and Temporal Analysis

As we have alluded to in previous chapters, environmental factors play a role in inpatient violence. Thus, we have analysed the spatial distribution of the incidents as well as their temporal distribution during a day (24 hours).

Figure 19 displays the frequency related to the specific location where the incidents have been reported. The highest number of incidents happened in the hallway (41.7%), followed by the bedroom (21.6%) and the dining room (9%). These findings are in line with what is reported in the literature. We can assume that a higher exposure to human interactions might lead to more episodes of escalation of violence and aggression. Environments such as the hallway, dining room and living room tend to be overcrowded and with lower levels of privacy. Both of these environmental factors have been considered as environmental precipitants of violence in psychiatry settings (Royal College of Psychiatrists, 1998). As mentioned earlier, it is always important to consider the phenomenon of violence as a combination of internal and external factors. The characteristics of the perpetrators and their inner state at the time of the aggression interact with the environment and its characteristics. The relationship with staff might also play a role. A higher frequency of aggression in certain environments might reflect the need to improve the rapport between the staff and patients.

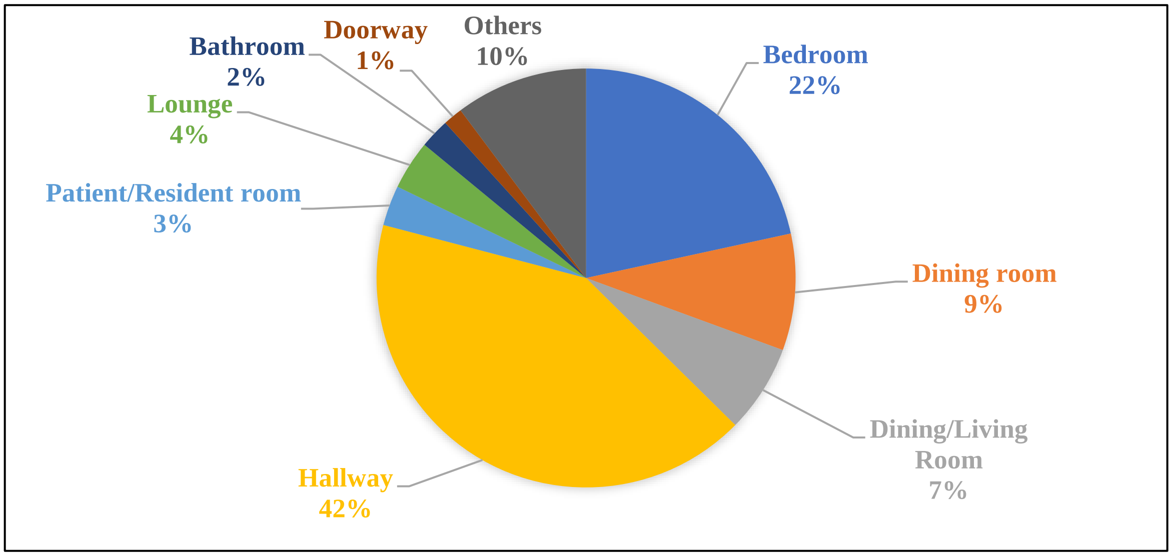


Figure 19. Specific location of incidents

An overview of how the aggression incidents are distributed within the entire arch of the day (24 hours) shows that most of the episodes happened in the afternoon, between 2 and 7 pm (Figure 20).

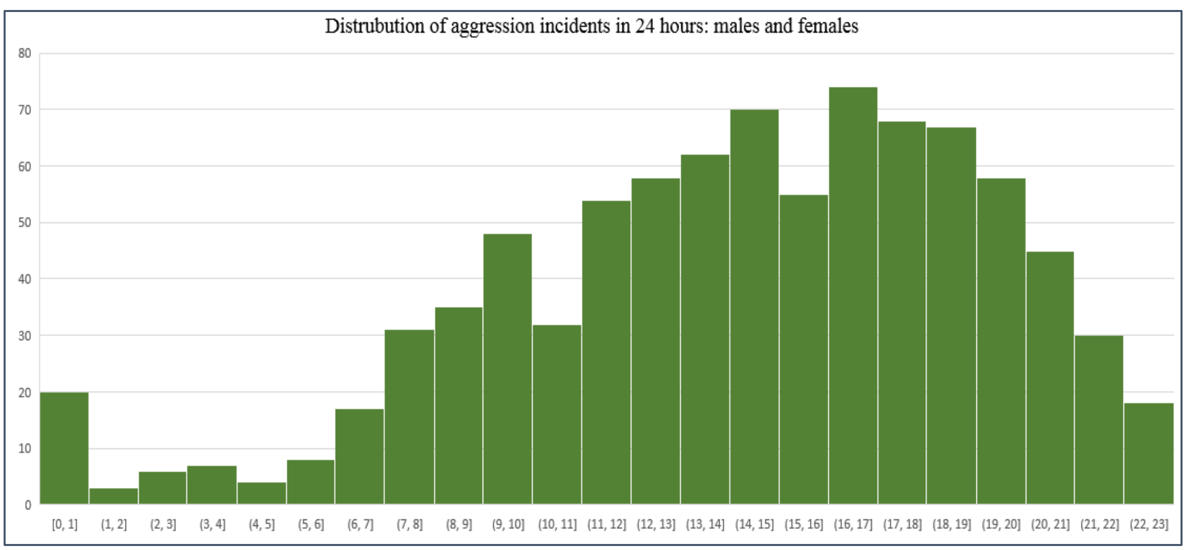


Figure 20. Distribution of aggression incident in 24 hours: males and females

By looking at the distribution of incidents over the 24 hours for males and females, we notice a few differences. The distribution of incidents committed by females has two peaks (Figure 22), between 12 am - 1pm and between 5 – 6 pm; the distribution of males (Figure 21) is more similar to the overall distribution presented above.

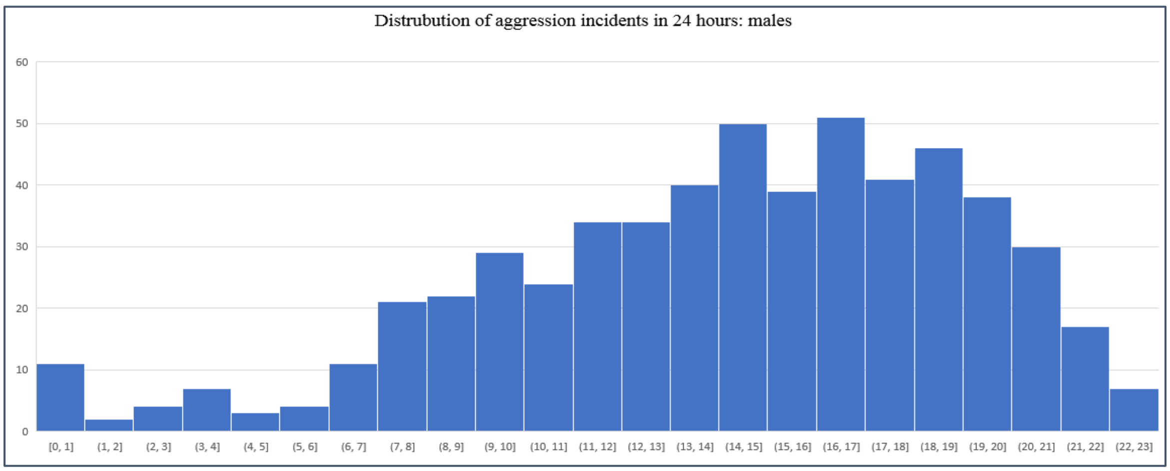


Figure 21. Distribution of aggression incident in 24 hours: males

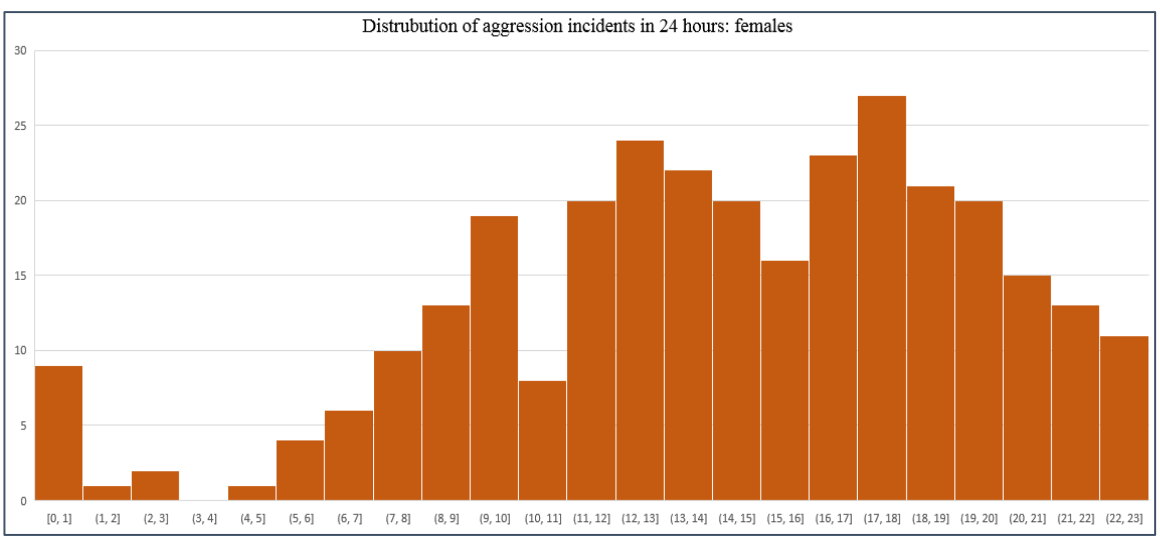


Figure 22. Distribution of aggression incident in 24 hours: females

As we can see in the bar graphs, the percentage of incidents that occur during night hours is very small, and it increases progressively starting from approximately 6 am. This general trend is intuitive: the risk is minimal during the night, and it increases during the day. For future research, it could be interesting to analyse the relationship between times of aggression, diagnosis and treatment plans.

4. Predictor for the classification of harm

One of the goals of our research is to investigate the risk factors most strongly associated with harmful incidents in the inpatient setting. For this purpose, we made a further categorization of the incidents reported by the SIRS. As previously showed, the original SIRS report classifies the incidents into 6 Levels:

1. Level 1-Near Miss
2. Level 2-No Harm Incident
3. Level 3-Harmful Incident - Mild Harm
4. Level 4-Harmful Incident - Moderate Harm
5. Level 5-Harmful Critical Incident – Severe Harm
6. Level 6-Harmful Critical Incident – Death

Based on this initial classification, we decided to divide the incidents into two main categories: incidents with no harm (Level 1 and Level 2) and incidents with harm (from Level 3 to Level 6: mild to critical harm). This classification will help us apply a logistic regression analysis on the data. Thus, the variable "Incident Severity for Reporting" became the dichotomic variable (dependent variable). Below are show the steps for re-coding this variable in R.

1. Level 1-Near Miss
2. Level 2-No Harm Incident

No Harm → Value "0" assigned in R

3. Level 3-Harmful Incident - Mild Harm
4. Level 4-Harmful Incident - Moderate Harm
5. Level 5-Harmful Critical Incident – Severe Harm
6. Level 6-Harmful Critical Incident – Death

Harm → Value "1" assigned in R

4.1 Risk of harm predictors

The predictors for our model have been chosen based on both the literature review and the results from our analysis in the previous sections. The last diagnosis received before the aggression incident is grouped into four (4) categories, that are most frequently associated to aggression episodes, according to our data as well as the literature review. As in the literature review, the history of violence has been also considered a possible factor in our model, and it has been split into two different variables. We also considered gender, age and the time between the admission to the psychiatric ward and the aggression episode.

		Estimate	Std. Error	z value	Pr(>|z|)
	Intercept	0.582394	0.416503	1.398	0.16203
	AGE	-0.027007	0.006647	-4.063	4.84e-05 ***
	GENDER	-0.129266	0.233563	-0.553	0.57995
History of violence	History of police intervention for violent behaviour	0.233833	0.073780	3.169	0.00153 **
	History of violence to others	-0.150915	0.063501	-2.377	0.01747 *
	Time to physical aggression after being admitted (in days)	0.003358	0.003367	0.997	0.31861
Last diagnosis received before the aggression incident (DSM-V)	Neurocognitive Disorders	-0.134072	0.401784	-0.334	0.73861
	Personality Disorders	0.388273	0.390827	0.993	0.32048
	Schizophrenia Spectrum and Other Psychotic Disorders	-0.960051	0.343833	-2.792	0.00524 **
	Substance-Related and Addictive Disorders	0.251586	0.515015	0.489	0.62519

*P < 0.05; **P < 0.01; *** P < 0.001

Table 14. Logistic regression. Risk of harm predictors

We find that the risk factors significantly associated to a harmful aggression are the age, history of police intervention for violent behaviour, history of violence to others, and a diagnosis within the “Schizophrenia Spectrum and Other Psychotic Disorders” (see Table 14). More specifically, the fact or most strongly associated to a harmful aggression is the age (p-value <0.001), followed by history of police intervention for violent behaviour (p-value <0.01), diagnosis of Schizophrenia spectrum and other psychotic disorders (p-value <0.01), and history of violence to others (p-value <0.05). These results are in line with the findings in literature. Different studies refer to the history of violence as a predictor of inpatient violence, and our results confirm that this element must be taken into serious consideration to take preventative measures for inpatient risk of harm.

The duration of admission did not seem to play a significant role in perpetrating harmful aggressions.

A correlation analysis was performed to confirm the relationship between age and the outcome variable (Figure 23).

```
                                                              Spearman's rank correlation rho

                                       Incident.Severity.for.Reporting PERSON.AGE
Incident.Severity.for.Reporting    cor 1                               -0.2256
                          N            616                             616
                     stat**                                            47747804
                    p-value                                            0.0000
---------------------------
                 PERSON.AGE    cor -0.2256                             1
                          N            616                             616
                     stat**            47747804
                    p-value            0.0000
---------------------------
       ** S
       HA: two.sided
```

Figure 23. Correlation analysis between age and Incident Severity for Reporting

With a p-value less than 0.001 and a coefficient of -0.2256, we confirmed a negative correlation between the two variables. Younger individuals are more at risk of committing harmful aggressions.

4.2 A Predictive Model

A further step in our research involves the realization of a possible predictive model capable of estimating the risk of harm associated with the aggression episodes reported in SIRS, followed by the evaluation of how efficiently this model works. The target variable is the same used before, “Incident Severity for Reporting”, which was split according to the 0/1 scheme following the criteria descripted above. To build and predict our model, we decided to focus on the patient population targeted by the same diagnostic categories used for our previous logistic regression analysis (Table 15). The process for the generation of a valid predictive model has three main phases

4.2.1 Phase 1: Data Partition

We partitioned our dataset in two sets: training and testing. The **training** represents the dataset utilized to build the predictive model, while the **testing** will be used as the validation set. In other words, the predictive model, built using the training dataset, will be evaluated using the testing dataset. In RStudio we set the outcome variable of interest and randomly split it in two sets (function: createDataPartition), according to the percentage indicated (p = 0.7) as in Figure 24.

```
Train <- createDataPartition(logistic$Incident.Severity.for.Reporting, p=0.7, list=FALSE)
training <- logistic[ Train, ]
testing <- logistic[ -Train, ]
summary(testing)
```

Figure 24. R Code: create data partition

4.2.2. Phase 2: Construction of Predication Model

We used a logistic regression to model "Incident Severity for Reporting" as a function of 5 predictors, selected based on our findings from both the literature review as well as the results of our preliminary analysis:

1. Police intervention for violent behaviour
2. Violence to others
3. Socially inappropriate disruptive behaviour
4. Age
5. Last diagnosis (most important) received before the aggression episode (four diagnoses)
    - Neurocognitive Disorders
    - Personality Disorders
    - Schizophrenia Spectrum and Other Psychotic Disorders
    - Substance-Related and Addictive Disorders

The model, called *mod_fit*, was initially applied to the training dataset (see Figure 25).

```
mod_fit <- train(Incident.Severity.for.Reporting ~      Police.Intervention..violent.behav.+
                 Violence.to.others+
                 Socially.inappropriate.disruptive.behaviour+

             Diagnosis..most.important....last.assessment.before.aggression + PERSON.AGE,
           data=training, method="glm", family="binomial")
```

dataset selected

Predictive model

Figure 25. R Code: predictive model using the training dataset

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5630  -0.7739  -0.5366   0.8265   2.5153

Coefficients:
                                                                                                     Estimate Std. Error z value Pr(>|z|)
(Intercept)                                                                                          1.177513   0.543563   2.166 0.030289 *
      Police.Intervention..violent.behav.1                                                           0.505122   0.401890   1.257 0.208802
      Police.Intervention..violent.behav.2                                                           0.206952   0.424888   0.487 0.626206
      Police.Intervention..violent.behav.3                                                           1.347121   0.526419   2.559 0.010497 *
      Police.Intervention..violent.behav.4                                                           0.844714   1.248835   0.676 0.498785
      Police.Intervention..violent.behav.5                                                           1.152022   0.550980   2.091 0.036541 *
      Violence.to.others1                                                                            0.597912   0.465544   1.284 0.199026
      Violence.to.others2                                                                           -0.142440   0.367255  -0.388 0.698127
      Violence.to.others3                                                                            0.272962   0.531353   0.514 0.607453
      Violence.to.others4                                                                           -0.570544   0.760072  -0.751 0.452867
      Violence.to.others5                                                                           -0.800403   0.456725  -1.752 0.079691 .
      Socially.inappropriate.disruptive.behaviour1                                                  -0.367306   0.489379  -0.751 0.452921
      Socially.inappropriate.disruptive.behaviour2                                                  -1.539155   0.511615  -3.008 0.002626 **
      Socially.inappropriate.disruptive.behaviour3                                                  -0.477822   0.406946  -1.174 0.240329
`Diagnosis..most.important....last.assessment.before.aggressionNeurocognitive Disorders`             0.060756   0.489312   0.124 0.901184
`Diagnosis..most.important....last.assessment.before.aggressionPersonality Disorders`                0.157282   0.484486   0.325 0.745457
`Diagnosis..most.important....last.assessment.before.aggressionSchizophrenia Spectrum and Other Psychotic Disorders` -1.525787   0.450028  -3.390 0.000698 ***
`Diagnosis..most.important....last.assessment.before.aggressionSubstance-Related and Addictive Disorders`  -0.752142   0.729888  -1.030 0.302781
PERSON.AGE                                                                                          -0.035240   0.008686  -4.057 4.97e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 502.66  on 431  degrees of freedom
Residual deviance: 423.68  on 413  degrees of freedom
AIC: 461.68

Number of Fisher Scoring iterations: 5
```

Figure 26. Logistic regression – predictive model applied to the training dataset

Police intervention for violent behaviour, Violence to others and Socially inappropriate disruptive behaviour have been transformed into factors. The numbers placed at the very right represent the factorial levels of those variables.

4.2.3 Phase 3: Testing and Validation of the Prediction Model

To evaluate how well our model predicts the target variable, we performed the following steps:

1. Applied the model to the testing set, which represents the observed values (function “predict”). This second model has been named *pred_mod.*

2. We performed a comparison between the outcome variable (“Incident Severity for Reporting”) in the new predictive model applied to the testing dataset (*pred_mod*) versus the real values of the training set (observed). For this purpose, we utilized the confusion matrix, a specific type of two-dimension contingency table capable of displaying the performance of a predictive algorithm. The positive class of the confusion matrix has been set up to the value “1”, corresponding to “harm” (see Figure 27).

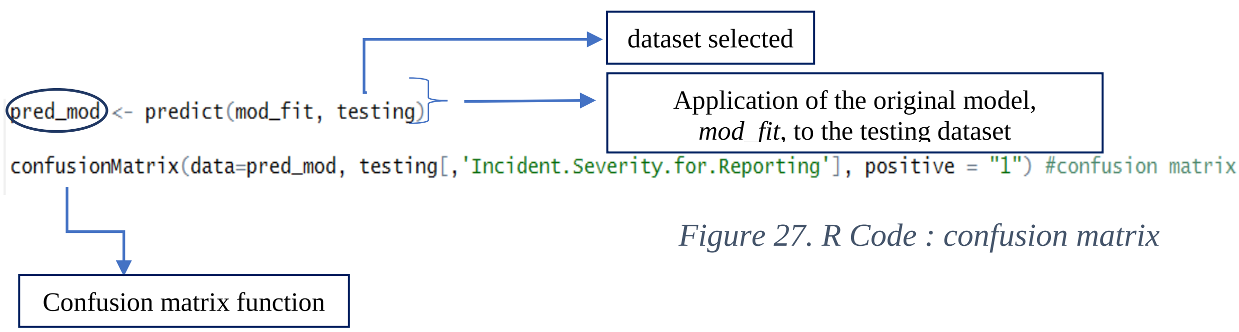


Figure 27. R Code : confusion matrix

Reference / Prediction	NO HARM (0)	HARM (1)
NO HARM (0)	True negatives	False negatives
HARM (1)	False positives	True positives

Figure 28. Confusion matrix

Figure 28 shows the theoretical structure of the expected confusion matrix. The columns represent the actual values, while the rows indicate the predictions:

- True negatives = values correctly identified as negatives (“0”)
- True positives = values correctly identified as positives (“1”)
- False positives = negatives; incorrectly identified as positives
- False negatives = positives; incorrectly identified as negatives

The term “confusion” refers to the fact that the model under examination is confused when it generates the predictions. This technique is very important because it gives information not only about the errors generated by the model, but also about what type of errors the model produces.

In Figure 29 we show the outcome of the confusion matrix for our model.

```
> confusionMatrix(data=pred_mod, testing[,'Incident.Severity.for.Reporting'], positive = "1") #confusion matrix
Confusion Matrix and Statistics

          Reference
Prediction   0   1
         0 124  35
         1  11  14

               Accuracy : 0.75
                 95% CI : (0.681, 0.8108)
    No Information Rate : 0.7337
    P-Value [Acc > NIR] : 0.342381

                  Kappa : 0.242
 Mcnemar's Test P-Value : 0.000696

            Sensitivity : 0.28571
            Specificity : 0.91852
         Pos Pred Value : 0.56000
         Neg Pred Value : 0.77987
             Prevalence : 0.26630
         Detection Rate : 0.07609
   Detection Prevalence : 0.13587
      Balanced Accuracy : 0.60212

       'Positive' Class : 1
```

Confusion matrix

Figure 29. Confusion matrix outcome with code in R

Accuracy	**(True positives + True negatives) / Total**	**75%**
Sensitivity	**Proportion of positives (“1”: harm) correctly identified**	**28.57%**
Specificity	**Proportion of negatives (“0”: no harm) correctly identified**	**91.85%**
Balanced Accuracy	(Sensitivity + Specificity) / 2	60.21%
Detection Rate	True positives / Total	0.07%
Positive Predictive Value (PPV)	True positives / (True positive + False positives)	56%
Negative Predictive Value (NPV)	True negatives / (True negatives + false negatives)	77.99%
Detection Prevalence	(True positive + False positives) / Total	13.59%
Prevalence	(False negatives + True positives) / Total = Total Positives / Total	26.63%

Table 15. Results confusion matrix

Table 15 describes the indicators obtained by building a confusion matrix to analyse the performance of our predictive model. The most significant results are described by the following three indicators: accuracy, sensitivity and specificity. The accuracy of our model is 0.75, which means that the model has predicted correctly 138 of the 184 episodes, with an accuracy of 75%. The investigation about what types of errors the model made reveals that 11 episodes of the 135 without any harm have been incorrectly predicted as at risk of harm, whereas 35 episodes of the 49 with some degree of harm have been incorrectly predicted as not at risk of harm. The specificity of the matrix indicates the percentages of no harm episodes correctly identified as a no risk of harm, amounting to 91.85% of cases. The sensitivity refers to the percentage of harmful incidents correctly identified as at risk of harm, which is equal to 28.57%.

4.2.4 Building the equation to predict the risk of harm, using the predictive model

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5630  -0.7739  -0.5366   0.8265   2.5153

Coefficients:
                                                                                                  Estimate Std. Error z value Pr(>|z|)
(Intercept)                                                                                       1.177513   0.543563   2.166 0.030289 *
      Police.Intervention..violent.behav.1                                                        0.505122   0.401890   1.257 0.208802
      Police.Intervention..violent.behav.2                                                        0.206952   0.424888   0.487 0.626206
      Police.Intervention..violent.behav.3                                                        1.347121   0.526419   2.559 0.010497 *
      Police.Intervention..violent.behav.4                                                        0.844714   1.248835   0.676 0.498785
      Police.Intervention..violent.behav.5                                                        1.152022   0.550980   2.091 0.036541 *
      Violence.to.others1                                                                         0.597912   0.465544   1.284 0.199026
      Violence.to.others2                                                                        -0.142440   0.367255  -0.388 0.698127
      Violence.to.others3                                                                         0.272962   0.531353   0.514 0.607453
      Violence.to.others4                                                                        -0.570544   0.760072  -0.751 0.452867
      Violence.to.others5                                                                        -0.800403   0.456725  -1.752 0.079691 .
      Socially.inappropriate.disruptive.behaviour1                                               -0.367306   0.489379  -0.751 0.452921
      Socially.inappropriate.disruptive.behaviour2                                               -1.539155   0.511615  -3.008 0.002626 **
      Socially.inappropriate.disruptive.behaviour3                                               -0.477822   0.406946  -1.174 0.240329
`Diagnosis..most.important....last.assessment.before.aggressionNeurocognitive Disorders`          0.060756   0.489312   0.124 0.901184
`Diagnosis..most.important....last.assessment.before.aggressionPersonality Disorders`             0.157282   0.484486   0.325 0.745457
`Diagnosis..most.important....last.assessment.before.aggressionSchizophrenia Spectrum and Other Psychotic Disorders` -1.525787 0.450028 -3.390 0.000698 ***
`Diagnosis..most.important....last.assessment.before.aggressionSubstance-Related and Addictive Disorders` -0.752142 0.729888 -1.030 0.302781
PERSON.AGE                                                                                       -0.035240   0.008686  -4.057 4.97e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 502.66  on 431  degrees of freedom
Residual deviance: 423.68  on 413  degrees of freedom
AIC: 461.68

Number of Fisher Scoring iterations: 5
```

Coefficients utilised for the linear predictor

The purpose of our model is to identify the probability of the risk of harm for a patient. The model is based on a logistic regression, described by the logit equation:

Logit equation: $$\operatorname{logit}(p) = \log\left(\frac{p}{1-p}\right) = \log(p) - \log(1-p) = -\log\left(\frac{1}{p} - 1\right).$$

Based on this equation, the estimated probability of the risk of harm can be calculated in the following way.

$$p_i = \frac{1}{1+e^{-\underbrace{(\beta_0+\beta_1 x_{1,i}+\cdots+\beta_k x_{k,i})}_{\text{Linear predictor}}}}.$$

P Risk of Harmful aggression = 1 / (1 + $e^{\text{- (linear predictor)}}$)

4.2.5 Example

The following example shows how this model can be used. Suppose we have the following RAI-MH of a patient:

Police intervention for violent behaviour	**2**
Violence to others	**2**
Socially inappropriate disruptive behaviour	**0**
Age	**20**
Last diagnosis (most important) received before the aggression episode (four possible diagnoses)	**Schizophrenia**

Table 16. Clinical example

Police intervention for violent behaviour: **2**

Socially inappropriate / disruptive behaviour: **0**

Linear predictor: 1.178 + 0.207 – 0.142 + 0 + 20*(-0.035) – 1.525 = – 0.982

Intercept

Violence to others: **2**

Age

Diagnosis

P Risk of Harmful aggression = 1 / 1 + $e^{\text{-(linear predictor)}}$ = 1 / (1 + $e^{-(-0.982)}$) = 0.727

The probability of risk of harmful aggression is P = 1 / (1 + $e^{-(\text{linear predictor})}$) =

$$1 / (1 + e^{-(-0.982)}) = \boxed{0.727}$$

If P > 0.5, then we classify the patient under the category risk of harm (category "1"). In our example, the probability is equal to 0.727, implying that the patient is likely to be at risk of harm (P > 0.5, category "1").

5. RHO CAP: performance among the residents who did not commit any aggression

In this section we look at the RHO CAP's performance among patients who did not commit any aggression. This analysis is needed to better understand the performance of the RHO CAP and its capacity to discriminate patients who are at risk of committing violence or aggression.

5.1 RHO CAP prediction power for risk of violence

From the patients' records (RAIs-MH) we considered only the ones related to new admissions and quarterly assessments (every 90 days), removing all the others (e.g. discharge assessments), for a total of 3452 charts. RHO CAPs related to discharge assessments cannot be used as the patient is discharged from the unit.

Table 17 shows the RHO CAP scores for admission and quarterly assessments from January 2016 to December 2017. About 73.55% of the RHO CAP scores refer to a low risk of harm to others. Considering that we are working with a patients' population that did not commit any act of aggression, we can conclude that the RHO CAP has demonstrated a good capacity in discriminating residents who were not at risk of harm to others or at risk of committing acts of aggression (regardless of the risk of harm). RHO CAP scores showing moderate and high risk of harm (26.45%) can be evaluated as false positives. It is important to note that the patients who received a medium or high RHO CAP score might represent a real risk in the inpatient setting, even if they did not commit any aggression.

		RHO CAP		**p-value > 0.05**
RAI-MH: Assessment Type	**0**	**1**	**2**	Total
Admission (total %) (% within admissions)	2098 (60.78%) (72.92%)	428 (12.4%) (14.88%)	351 (10.71%) (12.20%)	2877 (83.34%)
Quarterly (total %) (% within quarterly)	441 (12.78%) (76.7%)	80 (2.32%) (13.91%)	54 (1.56%) (9.39%)	575 (16.66%)
Total	2539 (73.55%)	508 (14.72%)	405 (11.73%)	3452

Table 17. Relationship between the RHO CAP and RAI-MHs for admission and quarterly assessments

	RHO CAP			p-value < 0.001
Diagnosis (DSM-V)	**0**	**1**	**2**	**Total**
Anxiety disorders (total %) (% within diagnostic category)	53 (1.54%) (88.33%)	3 (0.09%) (5%)	4 (0.12%) (6.67%)	60 (1.74%)
Bipolar and related disorders	331 (9.59%) (69.68%)	70 (2.03%) (14.74%)	74 (2.14%) (15.58%)	457 (13.76%)
Depressive disorders	582 (16.86%) (91.94%)	43 (1.25%) (6.79%)	8 (0.23%) (1.26%)	633 (18.34%)
Mood disorders	72 (2.09%) (79.12%)	14 (0.41%) (15.38%)	5 (0.14%) (5.49%)	91 (2.64%)
Neurocognitive disorders	101 (2.93%) (45.09%)	40 (1.16%) (17.86%)	83 (2.4%) (37.05%)	224 (6.49%)
Neurodevelopmental disorders	17 (0.49%) (50%)	8 (0.23%) (23.53%)	9 (0.26%) (26.47%)	34 (0.98%)
Personality disorders	128 (3.71%) (78.05%)	28 (0.81%) (17.07%)	8 (0.23%) 4.88%)	164 (4.75%)
Schizophrenia spectrum and other psychotic disorders	996 (28.85%) (69.94%)	253 (7.33%) (17.77%)	175 (5.07%) (12.29%)	1424 (41.25%)
Substance-related disorders	21 (0.61%) (75%)	2 (0.06%) (7.14%)	5 (0.14%) (17.86%)	28 (0.81%)
Trauma- and stressor-related disorders	18 (0.52%) (81.82%)	2 (0.06%) (9.09%)	2 (0.06%) (9.09%)	22 (0.64%)

	Large Sample Test Statistic	DF	p-value
Chi Squared	357.158	48	<0.001

Table 18. Relationship between Diagnosis and RHO CAP

The diagnoses associated ($\chi 2$ = 357.158, df = 48; p-value < 0.001) with the highest percentages of RHO CAP scores related to medium and high risk refer to the psychotic spectrum, neurocognitive disorders, neurodevelopmental disorders and bipolar and related disorders (Table 18). Neurocognitive disorders and neurodevelopmental disorders are the two diagnostic categories mostly associated with the RHO CAP scores predicting medium and high risk of harm to others, with 54.91% and 50%, respectively. However, the number of neurodevelopmental disorders represent only 0.98% of the sample (34 diagnoses), a number that is too small to be used as basis for analysis. On the other hand, neurocognitive disorders represent about 6.49% (224 diagnoses), the fourth most frequent diagnosis. It might be possible that the diagnostic category of neurocognitive disorders exposes the RHO CAP to a higher number of false positive predictions. It would be necessary to study, within the diagnosis of neurocognitive disorders, what are the factors associated with RHO CAP scores that flag medium and high risk.

The total of RHO CAP scores for patients who committed aggressions between 2016 and 2017 is reported in Table 19. The performance of the RHO CAP between the two samples of patients, aggressive and non-aggressive, is clearly different.

	RHO CAP last RAI-MH before incident			**Total**
	0	**1**	**2**	
Total	318 (36.55%)	202 (23.22%)	350 (40.23%)	870

Table 19. Total of RHO CAP scores for patient who committed aggression

We can conclude that RHO CAP showed a good performance in discriminating patients who represented a minimal or null risk to commit aggressions, without specifying the real risk of harm. In fact, the 73.55% of the RHO CAP scores, among patients who did not commit any inpatient aggression, resulted in the low risk category.

5.2 Probability of the occurrence of the first violent event

The time distribution of the first aggression incidents (2016-2017) that received a RHO CAP of 1 or 2 has been analysed. We calculated the amount of days between January 1st 2016 and each first episode of aggression for each patient. The first aggression occurred within an average of 323.59 days (SD = 216.16 days) from January 1st. The normality test showed that the distribution for the time to first occurrence is not normal (p-value = 4.83e-08). We therefor conducted further tests to find what distribution best fits our dataset. For this purpose, we used the ARENA Input Analyzer version 15. We found that the Beta distribution is the best fit for our data (Figure 30 and Table 20).

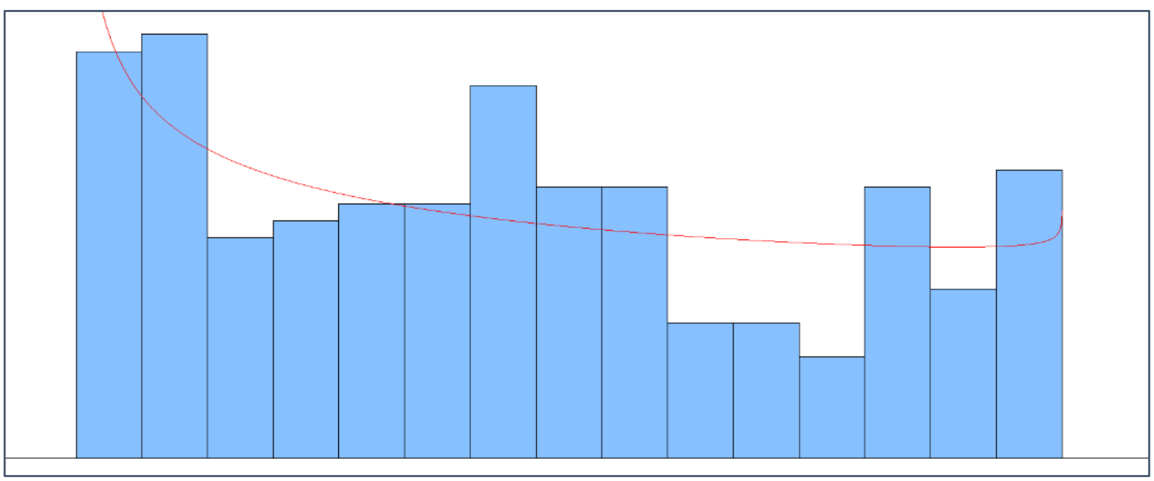


Figure 30. Beta distribution of first aggression incidents

Distribution: Beta	**Expression**: 5 + 721 * BETA(0.771, 0.973) **Square Error**: 0.005409
Chi Square Test	Number of intervals = 15 Degrees of freedom = 12 Test Statistic = 19 Corresponding p-value = 0.0908
Kolmogorov-Smirnov Test	Test Statistic = 0.0489 Corresponding p-value > 0.15
Data Summary	Number of Data Points = 225 Min Data Value = 5 Max Data Value = 726 Sample Mean = 324 Sample Std Dev = 216
Histogram Summary	Histogram Range = 5 to 726 Number of Intervals = 15

Table 20. Parameters for the best fit distribution for fist time of aggression occurrence

The Beta distribution is defined on the interval [0, 1] parametrized by two shape parameters, typically labelled with *p* and *q*, representing the lower and upper bounds.

$$\mathrm{B}(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx$$

The general formula of the probability density function of the beta distribution is:

.

$$f(x) = \frac{1}{\mathrm{B}(p,q)}\,\frac{(x-A)^{p-1}(B-x)^{q-1}}{(B-A)^{p+q-1}},\ \ A \le x \le B$$

In this equation A and B are the lower and upper bounds of the distribution. When A = 0 and B = 1 the distribution is called *standard beta distribution*:

$$f(x) = \frac{1}{\mathrm{B}(p,q)}\,x^{p-1}(1-x)^{q-1},\ \ 0 \le x \le 1$$

From table 1, we see that that the time to the first occurrence of aggression is:

5 + 721 * B (0.771, 0.973) in days.

5.3 Example

Suppose we would like to find out the probability that a newly admitted patient might commit an aggression within a certain time, say 150 days. To do so we solve for the

$$5 + 721 * \underbrace{Z\_BETA(0.771, 0.973)}_{\text{Z score for } \textit{first aggression}} = 150$$

Z_BETA(0.771, 0.973) = (150 – 5) / 721= 145 / 721 = 0.2

Thus, the probability of this patient committing an aggression within 150 days of admission is the cumulative probability for Beta (0,771, 0.973) that corresponds to a score of 0.2. This can be found by Excel using BETA_DIST (0.2,0/771,0.973,1) = 0.28 or a 28% chance that this patient will commit an aggression within 150 days of his or her stay.

6. Conclusion

Our data show that assaults have been committed predominantly by males between January 1st, 2016 and December 31st, 2017 (565 male aggressions vs 305 female aggressions). Among a total of 870 aggression, 648 (74.48%) were reported as not harmful. The analysis of the RHO CAP's behaviour revealed a significant relationship with different indicators including the last diagnosis before the aggression and the severity level of the episode. The overall accuracy of the RHO CAP was around 31.26%, displaying a higher accuracy for cases not at risk of harm (24.94%) than for cases at risk of harm (6.32%). Schizophrenia Spectrum and Other Psychotic Disorders and Neurocognitive Disorders are the two most frequent diagnostic categories (DSM-V) reported in the last assessment before the aggression episode, with 33.68% and 19.08% respectively. Analysing the harmful aggressions, we found that the major factors associated with these episodes were age, history of violence to others, police intervention for violent behaviour, and a diagnosis of psychosis. The predictive model developed for inpatient aggressions demonstrated an overall accuracy of 75%, with a specificity of 91.85% for episodes not at risk of harm, and a sensitivity of 28.57% for episodes at risk of harm. Among patients who didn't commit any aggression during the same time frame (2016 – 2017), the RHO CAP confirmed an overall accuracy of 73.55%. To support our interpretation of the RHO CAP performance among our sample of aggressions, we analysed the time distribution of the first aggression incidents (2016-2017) that received a RHO CAP of 1 or 2, accompanied by a practical example. The first aggression incidents' distribution is best represented by a Beta distribution, with an average of approximately 324 days (from January 1st, 2016) and a SD of 216 days.

Chapter VII

Conclusion

1. Summary of Major Findings

This thesis research confirms the importance of the RAI-MH as a comprehensive tool in mental health. The clinical measures and indicators implemented in this tool are used for a variety of purposes, supporting the transition towards a patient-centred healthcare system, where the allocation of resources is based on the utilization of services.

Given the scope of our research, we considered only some of the indicators implemented in the RAI-MH. The first sample, that focused on patients who committed aggressions, was predominantly composed by males (224 males and 113 females). The number of aggressions committed by males is almost double that committed by females (565 vs 305). Most of the incidents took place in the hallway, bedroom, dining room and living room, reaching a peak from 2 pm to 7 pm.

We found a significant relationship between the RHO CAP (the Risk of Harm to Others Clinical Assessment Protocol) and the following factors: severity of Incidents reported, education, last diagnosis received before the aggression episode, capacity of insight into mental health problem, and the aggressive behaviour scale. The performance of the RHO CAP has been studied by looking at the association with the Levels of Incident Severity, a measure reported by the clinical staff which indicates the degree of severity of the aggression episode. The RHO CAP showed an overall accuracy of approximately 31.26%. More specifically, the accuracy for the cases not at risk of harm was 24.94%, whereas the accuracy for the cases at risk of harm (mild, moderate or severe) was 6.32%. The most frequent diagnoses received before an episode of aggression refers to the following categories: Schizophrenia Spectrum and Other Psychotic Disorders (33.68%), Neurocognitive Disorders (19.08%), Personality Disorders (9.54%), and Bipolar and Related Disorders (5.75%). Neurocognitive Disorders showed the highest association (63.25%) with a RHO CAP score equal to 2 (high risk). Age (risk increases with younger individuals), history of violence to others, police intervention for violent behaviour, and a diagnosis referred to the psychotic spectrum, are the major factors that are associated with harmful aggressions in the inpatient setting.

Based on the data from both the literature review and our analyses, we developed a predictive model for inpatient harmful aggressions which considers the following factors: police intervention for violent behaviour, violence to others, socially inappropriate behaviour, age, and the last diagnosis received before the aggression episode among four main diagnostic categories. The model demonstrated an overall accuracy of 75%, with a specificity to identify episodes considered not at risk of harm equal to 91.85%, and a sensitivity to identify episodes considered harmful equal to 28.57%. The main purpose that led us to generate this algorithm was to provide more information and insights in predicting harm in inpatient setting. Some of the constructs that emerged as strong predictors warrant further studies in the future and have the potential to create different models that can serve as reliable “red flags” for clinicians.

We also analysed the RHO CAP performance among patients who did not commit any aggression during the time frame considered (January 2016 – December 2017), reviewing retrospectively patient records related to admission and quarterly assessments. The RHO CAP demonstrated an overall accuracy of 73.55% in discriminating patients who represented a minimal or null risk to commit aggressions. This result stresses the usefulness of this indicator.

This is the first study that analysed the RHO CAP performance within two patients’ populations: patients who did not commit any aggression and patients who committed aggression. In the first case we evaluated its capacity of predicting the risk of harm within the aggressive patients, in the second case we evaluated its capacity to predict the risk of aggressive behaviours.

This distinction is important, due to the complexity of the concept of aggression. In chapter one we gave an overview of the two types of aggression: impulsive and instrumental. The most important difference lies in the intent to harm someone, which characterizes the impulsive type. Impulsive aggression is also characterized by impulsive behaviour and anger (Bushman & Anderson, 2001), whereas instrumental (or premeditated) aggression is the result of a mental plan, a strategy to obtain something. The intent of harming someone, according to this theoretical framework, results mostly from a reduced capacity to control the impulses as well as the difficulty of considering the consequences of the actions. The other theoretical framework explored in Chapter 1 relates to the dichotomy reactive/proactive aggression (Cima et al., 2013; Rosell & Siever, 2017), where reactive aggression represents the counterpart of the impulsive aggression, driven by frustration and the tendency to interpret stimuli perceived as a threat. As explained by different authors these categories, in both the dichotomies, are not rigid, and the distinction of one type of aggression from another is not always possible. However, they can be very helpful in understanding the basic principles underlying aggressive behaviours.

In our study, the RHO CAP demonstrated to be a good predictive measure for screening the risk of inpatient aggressive behaviours, however, it showed a significantly reduced accuracy in detecting the real risk of harm of a resident. This raises some important questions: is there a relationship between the risk of harm and the intent of harm, defined as a core component of the impulsive aggression? Is the RHO CAP an indicator that is more useful to flag the general risk of possible aggressive behaviours, rather than a flag for a specific risk of harm? More research is needed to address these questions by including, as possible predictors of harm to others, indicators capable of exploring the concepts of impulsivity, frustration, cognitive distortion related to an exaggerated interpretation of threat stimuli, and anger.

2. Limitations and future directions

While we had generous access to data from our hospital partners, our research still faced some limitations. We analysed the phenomenon of aggression from the point of view of clinicians. We did not have data on the point of view of patients. As alluded to in the literature, the patient side of the equation is important but it has received little attention. It would be useful to augment our analysis with patient survey data that reflects their perspective on how they perceive the staff's behaviours and the environment.

Another limitation is in time period we have consider, from 2016 to 2017. Despite the significant amount of data that we have collected, longer observation times would help in filling some of the missing data (such as that on the level of education or diagnosis).

Studying attitude and frustration of the clinical staff could be another strategy to provide insights about personal and interpersonal dynamics that are involved in inpatient units, and their relationship with episodes of aggression and violence. It might be interesting to study how specific disorders impact on the staff's frustration and behaviour, and how, in turn, this interaction might affect the number and severity of aggressions in inpatient units.

It is always useful to keep in mind that the accuracy as well as the subjectivity of the clinical data gathered might represent a limitation. The engagement of the clinical staff in using the RAI-MH might represent an important variable to consider. For example, if this tool is relatively new to a clinician or if workload and stress are very intense within the unit, the accuracy in completing the assessment might be affected.

The type of aggression was not addressed in our study. Amore et al. (2008) differentiated between three types of aggression: verbal, physical and against an object. Further analysis can be conducted to investigate the causal relationship between the aggressions and the specific types of incident. This research direction might be useful not only for a better understanding of the phenomenon, but also to provide insights about the environmental design of a psychiatric ward. There might be objects that patients tend to use more than others, or locations that pose more risk than others.

We also noted that many diagnoses changed over the admission period. A future research direction might focus on studying what diagnoses tend to change more frequently, what are the factors associated to this phenomenon and if it has any implication on inpatient violence.

Some studies reported that the duration of admission is a factor associated with inpatient aggression (Chou, Lu, & Mao, 2002), however, we did not find a significant association between the duration of admission and harm to others. Further studies might try to clarify this point as well. If there will be solid evidences proving that the duration of admission is a factor contributing to inpatient violence, then some strategies could be tailored to mitigate this effect, such as granting weekend passes.

Some authors have referred to involuntary admission as a factor associated with violent or aggressive behaviour (Johnson, 2004; Serper et al., 2005; Woods and Ashley, 2007; Soliman & Reza, 2001). However, this is an aspect that did not receive enough attention, relative to other factors. More efforts in this direction would help us understand to what extent the nature of admission might contribute to inpatient violence. The RAI-MH provides information about the type of admission, together with other meaningful sociodemographic information, confirming its critical importance not only for assessment and care planning purposes, but also for research purposes.

Bibliography

Abderhalden, C., Hahn, S., Bonner, Y. D. B. & Galeazzi, G. M. (2006). Users' perceptions and views on violence and coercion in mental health care. In: D. Ritcher & R. Whittington (Eds). Violence in Mental Health Settings: Causes, Consequences and Management. (pp. 69–92). New York: Springer

Allen, J. J., & Anderson, C. A. (2017). General Aggression Model. In P. Roessler, C. A. Hoffner, & L. van Zoonen (Eds.) International Encyclopedia of Media Effects. WileyBlackwell. DOI: 10.1002/9781118783764.wbieme0078

Allen, J. J., Anderson, C. A., & Bushman, B. J. (2018). The General Aggression Model. Current Opinion in Psychology, 19, 75–80. https://doi.org/https://doi.org/10.1016/j.copsyc.2017.03.034

American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 4th ed. Washington (DC): APA; 1994.

Amore, M., Menchetti, M., Tonti, C., Scarlatti, F., Lundgren, E., Esposito, W. and Berardi, D. (2008). Predictors of violent behavior among acute psychiatric patients: Clinical study. Psychiatry and Clinical Neurosciences, 62: 247–255. doi:10.1111/j.1440-1819.2008.01790.x

Anderson T.R., Bell C.C., Powell T.E., et al. (2004). Assessing psychiatric patients for violence. Community Mental Health Journal 40, 379–399.

Anderson, C. A., & Bushman, B. J. (2002). Human Aggression. Annual Review of Psychology, 53(1), 27–51. https://doi.org/10.1146/annurev.psych.53.100901.135231

Anderson, C.A., Bushman, B.J., (2002). Human aggression. Annual Review of Psychology 53, 27–51.

Andrews, D. A., & Bonta, J. (1995). Level of Service Inventory—Revised. Toronto: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The Recent Past and Near Future of Risk and/or Need Assessment. Crime & Delinquency, 52(1), 7–27. https://doi.org/10.1177/0011128705281756

Arango, C., Barba, A. C., Gonzalez-Salvador, T., & Ordonez, A. C. (1999). Violence in inpatients with schizophrenia: A prospective study. Schizophrenia Bulletin, 25, 493-503.

Arnetz, J. E. & Arnetz, B. B. (2001). Violence towards health care staff and possible effects on the quality of patient care. Social Science and Medicine, 52, 417–427.

Bandura, A. (1969). Principles of behavior modification. New York: Holt, Rinehart & Winston.

Bandura, A. (1973). Aggression: A social learning analysis. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1979). Psychological mechanisms of aggression. In M. VonCranach, K. Foppa, W. LePenies, & D. Ploog (Eds.), Human ethology: Claims and limits of a new discipline. Cambridge: Cambridge University Press.

Bandura, A. (2001). Social cognitive theory: an agentic perspective. Annu. Rev. Psychol. 52:1–26

Barkataki, I., Kumari, V., Das, M., Taylor, P., & Sharma, T. (2006). Volumetric structural brain abnormalities in men with schizophrenia or antisocial personality disorder. Behavioural Brain Research, 169(2), 239–247. https://doi.org/https://doi.org/10.1016/j.bbr.2006.01.009

Barlow, K., Grenyer, B., & Ilkiw-Lavalle, O. (2001). Prevalence and precipitants of aggression in psychiatric inpatient units. The Australian and New Zealand journal of psychiatry (Vol. 34). https://doi.org/10.1080/000486700271

Berk, R., Sherman, L., Barnes, G., Kurtz, E., & Ahlman, L. (2009). Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning. Journal of the Royal Statistical Society Series A (Vol. 172). https://doi.org/10.1111/j.1467-985X.2008.00556.x

Berkowitz L. (1989). Frustration-aggression hypothesis: examination and reformulation. Psychol. Bull. 106:59–73

Berkowitz L. 1990. On the formation and regulation of anger and aggression: a cognitive-neoassociationistic analysis. Am. Psychol. 45:494–503

Berkowitz L. (1993). Pain and aggression: some findings and implications. Motiv. Emot. 17:277–93.

Björkdahl, A., Palmstierna, T., & Hansebo, G. (2010). The bulldozer and the ballet dancer: Aspects of nurses' caring approaches in acute psychiatric intensive care. Journal of psychiatric and mental health nursing. 17. 510-8. 10.1111/j.1365-2850.2010.01548.x.

Bjørkly, Stål. (2006). Psychological theories of aggression: Principles and application to practice. Violence in Mental Health Settings: Causes, Consequences, Management. 27-46. 10.1007/978-0-387-33965-8_2

Blair, R. J. R. (2004). The roles of orbital frontal cortex in the modulation of antisocial behavior. Brain and Cognition, 55(1), 198–208. https://doi.org/https://doi.org/10.1016/S0278-2626(03)00276-8

Bonta, J., Harman, W. G., Hann, R. G., & Cormier, R. B. (1996). The prediction of recidivism among federally sentenced offenders: A revalidation of the SIR scale. Canadian Journal of Criminology, 38, 61–79. Retrieved from http://www.ccja-acjp.ca/en/cjc.html

Bos, J.T., D.H. Frijters, C. Wagner, G.I. Carpenter et al. (2007), "Variations in quality of Home Care between sites across Europe, as measured by Home Care Quality Indicators", Aging Clinical and Experimental Research, Vol. 19, No. 4, pp. 323-329, August.

Botz, C. K., Bestard, S., Demaray, M., & Molloy, G. (1993). Resource Utilization Groups (RUGs): Defining Chronic Care, Rehabilitation and Nursing Home Case Mix in Canada. Healthcare Management Forum, 6(4), 5–11. https://doi.org/10.1016/S0840-4704(10)61129-5

Bowers L, Simpson A, Alexander J, et al. (2005). The nature and purpose of acute psychiatric wards: the Tompkins Acute Ward Study. J Mental Health; 14:625–635.

Bowers L. (2002) Dangerous and Severe Personality Disorder. Reaction and Role of the Psychiatric Team. Routledge, London.

Bowers, L. (2007). The City-128 study of observation and outcomes. BMC Psychiatry 2, 1, 128. https://doi.org/10.1186/1471-244X-7-S1-S122

Bowers, L., Ross, J., Nijman, H., Muir- Cochrane, E., Noorthoorn, E., & Stewart, D. (2012). The scope for replacing seclusion with time out in acute inpatient psychiatry in england. Journal of Advanced Nursing, 68(4), 826-835. 10.1111/j.1365-2648.2011.05784.x

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984). Classification and regression trees. Wadsworth and Brooks/Cole, Monterey, CA

Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brown RL, Rounds LA. (1995). Conjoint screening questionnaires for alcohol and other drug abuse: criterion validity in a primary care practice. Wis Med J. ;94(3):135–140.

Burrows AB, Morris JN, Simon SE, et al. (2000). Development of a minimum data set-based depression rating scale for use in nursing homes. Age Ageing.;29(2):165–172.

Bushman, B. J., Anderson, C. A. (2001). Is it time to pull the plug on the hostile versus instrumental aggression dichotomy? Psychological Review, 108(1), 273-279. 10.1037/0033-295X.108.1.273

Carpenter, G.I., N. Ikegami, G. Ljunngren, E. Carrillo and B.E. Fries (1997). “RUG-III and Resource Allocation: Comparing the Relationship of Direct Care Rime with Patient Characteristics in Five Countries”, Age and Ageing, Vol. 26-S2, pp. 61-65.

Carpenter, I. and Hirdes, J.P. (2013) Using interRAI Assessment Systems to Measure and Maintain Quality of Long-Term Care. In: A Good Life in Old Age? Monitoring and Improving Quality in Long-Term Care, OECD/European Commission 2013. https://doi.org/10.1787/9789264194564-7-en

Catherine Hawes, Brant E. Fries, Mary L. James, Marylou Guihan (2007). Prospects and Pitfalls: Use of the RAI-HC Assessment by the Department of Veterans Affairs for Home Care Clients, The Gerontologist, Volume 47, Issue 3, 1 June 2007, Pages 378–387, https://doi.org/10.1093/geront/47.3.378

Caulkins, J., Cohen, J., Gorr, W., & Wei, J. (1996). Predicting criminal recidivism: A comparison of neural network models with statistical methods. Journal of Criminal Justice, 24(3), 227–240. https://doi.org/https://doi.org/10.1016/0047-2352(96)00012-8

Chou K.R., Lu R.B. & Mao W.C. (2002) Factors relevant to patient assaultive behavior and assault in acute inpatient psychiatric units in Taiwan. Archives of Psychiatric Nursing 16, 187–195.

Chou, K., Chi, I., & Leung, J. (2010). Applying Resource Utilization Groups (RUG-III) in Hong Kong Nursing Homes. Canadian Journal on Aging / La Revue Canadienne Du Vieillissement, 27(3), 233-239. doi:10.3138/cja.27.3.233.

Cima, M., Raine, A., Meesters, C., & Popma, A. (2013). Validation of the dutch reactive proactive questionnaire (RPQ): Differential correlates of reactive and proactive aggression from childhood to adulthood. Aggressive Behavior, 39(2), 99-113. 10.1002/ab.21458

Clauser, S. B., & Fries, B. E. (1992). Nursing home resident assessment and case-mix classification: Cross-national perspectives. Health Care Financing Review, 13(4), 135–155.

CMS.gov (2017). Centers for Medicare & Medicaid Services. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ProspMedicareFeeSvcPmtGen/index.html. Accessed 12 Jan 2018.

Cowin L, Davies R, Estall G , Berlin T, Fitzgerald M, HootS (2003). De-escalating aggression and violence in the mental health setting. International Journal of Mental Health Nursing 2003;12(1):64–73. [PUBMED: 14685961]

Cutcliffe, J. R., & Riahi, S. (2013a). Systemic perspective of violence and aggression in mental health care : Towards a more comprehensive understanding and conceptualization : Part 1, 558–567. https://doi.org/10.1111/inm.12029

Cutcliffe, J. R., & Riahi, S. (2013b). Systemic perspective of violence and aggression in mental health care : Towards a more comprehensive understanding and conceptualization : Part 2, 568–578. https://doi.org/10.1111/inm.12028

Dack, C., Ross, J., Papadopoulos, C., Stewart, D., & Bowers, L. (2013). A review and meta- analysis of the patient factors associated with psychiatric in- patient aggression. Acta Psychiatrica Scandinavica, 127(4), 255–268. https://doi.org/10.1111/acps.12053

Devriendt, E, Wellen, NI, Flamaing, J. (2013).The interRAI Acute Care Instrument incorporated in an eHealth system for standardized and web-based geriatric assessment: strengths, weaknesses, opportunities and threats in the acute hospital setting. BMC Geriatr. 2013;13:90. doi:10.1186/1471-2318-13-90.

Distributions. Journal of Statistical Software, 64(4), 1-34. URL http://www.jstatsoft.org/v64/i04/.

DoH (2005). Department of Health. Mental Capacity Act 2005. London:HMSO, 2005.

Dollard, J., Doob, L., Miller, N., Mowrer, O. H., & Sears, R. R. (1939). Frustration and aggression. New Haven, Conn: Yale University Press.

Douglas, K. S., Guy, L. S., & Hart, S. D. (2009). Psychosis as a Risk Factor for Violence to Others: A Meta-Analysis, 135(5), 679–706. https://doi.org/10.1037/a0016311

Duhart, D. T. (2001). Bureau of Justice Statistics Violence in the Workplace , 1993-99. BJS Statistician, December 2. Retrieved from https://www.bjs.gov/content/pub/pdf/vw99.pdf

Duxbury, J. (2002). An evaluation of staff and patient views of and strategies employed to manage inpatient aggression and violence on one mental health unit: A pluralistic design. Journal of Psychiatric and Mental Health Nursing, 9, 325–337.

Duxbury, J., & Whittington, R. (2005). Causes and management of patient aggression and violence: staff and patient perspectives. Journal of Advanced Nursing, 50, 469–478.

E.Y. Tjam, G.A. Heckman, S. Smith, et al. (2012). Predicting heart failure mortality in frail seniors: comparing the NYHA functional classification with the Resident Assessment Instrument (RAI) 2.0. Int. J. Cardiol., 155 (2012), pp. 75-80.

Ellen Dellefield, Mary. (2006). Using the Resource Utilization Groups (RUG-III) System as a Staffing Tool in Nursing Homes. Geriatric nursing (New York, N.Y.). 27. 160-5. 10.1016/j.gerinurse.2006.02.004.

Ewing JA. (1984). Detecting Alcoholism. The CAGE Questionnaire. JAMA.;252(14):1905–1907. doi:10.1001/jama.1984.03350140051025

Folstein, M.F., S.E. Folstein and P.R. McHugh (1975). “Mini-mental State. A Practical Method for Grading the Cognitive State of Patients for the Clinician”, Journal of Psychiatric Research, Vol. 12, No. 3, pp. 189-198.

Foster, C., Bowers, L. & Nijman, H. (2007). Aggressive behaviour on acute psychiatric wards: Prevalence, severity and management. Journal of Advanced Nursing, 58 (2), 140–149.

Freud, S. (1920). Beyond the pleasure principle. Standard Edition, (Vol. 18). London: Hogarth Press.

Fries BE, Simon SE, Morris JN, et al. (2001). Pain in US nursing homes: validating a pain scale for the minimum data set. Gerontologist. 41(2):173–179.

Fries, B.E., C.J. Fahey, C. Hawes, B.C. Vladeck, J. Morris, C. Phillips, H. Fredeking, J. Hirdes, D.G. Sinclair, J.T. King et al. (2003). “Implementing the Resident Assessment Instrument: Case Studies of Policymaking for Long-term Care in Eight Countries”, Milbank Memorial Fund, p. 129.

Fries, B.E., Schneider, D.P., Foley, W.J., Gavazzi, M., Burke, R., & Cornelius, E. (1994). Refining a case mix measure for nursing homes: Resource Utilization Groups (RUG-III). Medical Care, 32, 668–685.

Gardner, K. J., Archer, J., & Jackson, S. (2012). Does maladaptive coping mediate the relationship between borderline personality traits and reactive and proactive aggression? Aggressive Behavior, 38(5), 403-413. 10.1002/ab.21437

Gottfredson, S. D., & Gottfredson, D. M. (1986). Accuracy of prediction models. In A. Blumstein, J. Cohen, J. Roth,&C. A. Visher (Eds.), Criminal careers and “career criminals” (Vol. 2, pp. 212-290). Washington, DC: National Academy of Sciences Press.

Gottfredson, S. D., & Moriarty, L. J. (2006). Statistical Risk Assessment: Old Problems and New Applications. Crime & Delinquency, 52(1), 178–200. https://doi.org/10.1177/0011128705281748

Haglund, K; von Knorring, L; von Essen, L. (2006). Psychiatric wards with locked doors – advantages and disadvantages according to nurses and mental health nurse assistants. Journal of Clinical Nursing, 15(4), 387–394.

Hahn, S., Muller, M., Needham, I., Dassen, T., Kok, G. & Halfens, R. (2010). Factors associated with patient and visitor violence experienced by nurses in general hospitals in

Switzerland: A cross-sectional survey. Journal of Clinical Nursing, 19 (23–24), 3535–3546.

Haller, J. (2018). The role of central and medial amygdala in normal and abnormal aggression: A review of classical approaches. Neuroscience & Biobehavioral Reviews, 85, 34–43. https://doi.org/https://doi.org/10.1016/j.neubiorev.2017.09.017

Hamilton, Z., Neuilly, M.-A., Lee, S., & Barnoski, R. (2014). Isolating modeling effects in offender risk assessment. Journal of Experimental Criminology (Vol. 11). https://doi.org/10.1007/s11292-014-9221-8

Harper, P. R. (2005). A review and comparison of classification algorithms for medical decision making. Health Policy, 71(3), 315–331. https://doi.org/https://doi.org/10.1016/j.healthpol.2004.05.002

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. Criminal Justice and Behavior, 20, 315–335. doi:10.1177/0093854893020004001

Hartvig, P., Alfarnes, S., Østberg, B., Skjønberg, M., & Moger, T. A. (2006). Brief checklists for assessing violence risk among patients discharged from acute psychiatric facilities: A preliminary study. Nordic Journal of Psychiatry, 60(3), 243–248. https://doi.org/10.1080/08039480600780532

Hiday VA, Swartz MS, Swanson JW, Borum R, Wagner HR. (1998). Male-female differences in the setting and construction of aggression among people with severe mental illness. Soc Psychiatry Psychiatr Epidemiol 1998; 33:S68-S74.

Hirdes JP, Curtin-Telegdi N, Mathias K, et al. (2011). InterRAI Mental Health Clinical Assessment Protocols (CAPs) for Use with Community and Hospital-Based Mental Health Assessment Instruments. Version 9.1. Washington, DC: interRAI.

Hirdes JP, Fries BE, Morris JN et al. (1999). Integrated health information systems based on the RAI/MDS series of instruments. Healthcare Manage Forum; 12:30–40.

Hirdes JP, Ljunggren G, Morris JN, et al. (2008). Reliability of the interRAI suite of assessment instruments: a 12-country study of an integrated health information system. BMC Health Serv Res.; 8:277.

Hirdes JP, Marhaba M, Smith TF, et al. (2000). Development of the Resident Assessment Instrument-Mental Health (RAI-MH). Hospital Quarterly; 4(2):44–51.

Hirdes, J. P., Smith, T. F., Rabinowitz, T., Yamauchi, K., Pérez, E., Telegdi, N. C., … Fries, B. E. (2002). The resident assessment instrument-mental health (RAI-MH): Inter-rater reliability and convergent validity. The Journal of Behavioral Health Services & Research, 29(4), 419–432. https://doi.org/10.1007/BF02287348

Hirdes, J.P., B.E. Fries, J.N. Morris, K. Steel, V. Mor, D. Frijters, S. LaBine, C. Schalm, M.J. Stones, G. Teare et al. (1999). "Integrated Health Information Systems Based on the RAI/MDS Series of Instruments", Health Manage Forum, Vol. 12, No. 4, pp. 30-40.

Hirdes, J.P., Ljunggren, G., Morris, J.N., et al. (2008). Reliability of the interRAI suite of assessment instruments: A 12-country study of an integrated health information system. BMC Health Services Research, 8, 277.

Huesmann LR. (1986). Psychological processes promoting the relation between exposure to media violence and aggressive behavior by the viewer. J. Soc. Issues 42:125–40

Huesmann LR. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. See Geen & Donnerstein 1998, pp. 73–109

Institute of Medicine (1986). Improving the quality of care in nursing homes. Washington, DC: National Academy Press.

InterRAI (2018). Case-Mix Classification. www.interrai.org/classification.html Accessed Jan 14th 2018

Iozzino, L., Ferrari, C., Large, M., Nielssen, O., & De, G. (2015). Prevalence and Risk Factors of Violence by Psychiatric Acute Inpatients: A Systematic Review and Meta-Analysis. PLoS ONE, 10(6). https://doi.org/10.1371/journal.pone.0128536

J. Gindin, A. Walter-Ginzburg, M. Geitzen, et al. (2007). Predictors of rehabilitation outcomes: A comparison of Israeli and Italian geriatric post-acute care (PAC) facilities using the minimum data set (MDS). J Am Med Dir Assoc, 8 (2007), pp. 233-242.

Jenkins, E. J. & Bell, C. C. (1997). Exposure and response to community violence among children and adolescents. In J. Osofsky (Ed.), Children a Violent Society. New York: Guilford Press.

Johnson A, (2003). Caregiver's corner webinars: crisis situations & de-escalation strategies. http://hdsa.org/living-with-hd/caregivers-corner-webinars/crisis-situations-de-escalation-strategies/ (accessed 3 February 2015).

Johnson M.E. (2004). Violence on inpatient psychiatric units: state of the science. Journal of the American Psychiatric Nurses Association 10, 113–121.

Kumari, V., Uddin, S., Premkumar, P., Young, S., Gudjonsson, G. H., Raghuvanshi, S., … Das, M. (2013). Lower anterior cingulate volume in seriously violent men with antisocial personality disorder or schizophrenia and a history of childhood abuse. Australian & New Zealand Journal of Psychiatry, 48(2), 153–161. https://doi.org/10.1177/0004867413512690

LeBel, J. & Goldstein, R. (2005). Special section on seclusion and restraint: The economic cost of using restraint and the value added by restraint reduction or elimination. Psychiatric Services, 56 (9), 1109–1114.

Leutgeb, V., Leitner, M., Wabnegger, A., Klug, D., Scharmüller, W., Zussner, T., & Schienle, A. (2015). Brain abnormalities in high-risk violent offenders and their association with psychopathic traits and criminal recidivism. Neuroscience, 308, 194–201. https://doi.org/https://doi.org/10.1016/j.neuroscience.2015.09.011

Littrell, K. H.; Littrell, S. (1998). Current understanding of violence and aggression: Assessment and treatment. Journal of Psychosocial Nursing & Mental Health Services, 36(12), 18–24.

Liu, J. (2004). Concept analysis: Aggression. Issues in Mental Health Nursing, 25(7), 693–714. https://doi.org/doi.org/10.1080/01612840490486755

Liu, Y. Y., Yang, M., Ramsay, M., Li, X. S., & Coid, J. W. (2011). A Comparison of Logistic Regression, Classification and Regression Tree, and Neural Networks Models in Predicting Violent Re-Offending. Journal of Quantitative Criminology, 27(4), 547–573. https://doi.org/10.1007/s10940-011-9137-7

Lobbestael, J., Cima, M., & Arntz, A. (2013). The Relationship Between Adult Reactive and Proactive Aggression, Hostile Interpretation Bias, and Antisocial Personality Disorder. Journal of Personality Disorders, 27(1), 53–66. https://doi.org/10.1521/pedi.2013.27.1.53

Luckhoff, M., Jordaan, E., Swart, Y., Cloete, K. J., Koen, L. & Niehaus, D. J. H. (2013). Retrospective review of trends in assaults and seclusion at an acute psychiatric ward over a 5-year period. Journal of Psychiatric and Mental Health Nursing, 20 (8), 687–695.

Marangos-Frost S. & Wells D. (2000). Psychiatric nurses' thoughts and feelings about restraint use: a decision dilemma. Journal of Advanced Nursing 31, 362–369.

Marie Laure Delignette-Muller, Christophe Dutang (2015). fitdistrplus: An R Package for Fitting

Martin, L., & Hirdes, J. P. (2009). Mental Health Needs and Service Use in Ontario. Healthcare Management Forum, 22(1), 40–46. https://doi.org/10.1016/S0840-4704(10)60291-8

Martin, L., Perlman, C., & Bieling, P. (2015). Using the RAI-MH to support shared decision-making in mental healthcare. Healthcare Management Forum, 28(4), 163–166. https://doi.org/10.1177/0840470415581231

McDermott, B. E., & Holoyda, B. J. (2014). Assessment of aggression in inpatient settings. CNS Spectrums, 19, 425–431. https://doi.org/10.1017/S1092852914000224

Meehan, T., Vermeer, C. & Windsor, C. (2000). Patients' perceptions of seclusion: A qualitative investigation. Journal of Advanced Nursing, 31, 370–377.

Merriam-Webster.com. Retrieved February 2nd, 2018, from https://www.merriam-webster.com/dictionary/violence.

Michie, S. & West, M. A. (2004). Managing people and performance: An evidence based framework applied to health service organizations. International Journal of Management Reviews, 5 (6), 91–111.

Morris JN, Fries BE, Mehr DR, et al. (1994). MDS Cognitive Performance Scale. J Gerontol. 1994;49(4):M174–M182.

Morris JN, Fries BE, Morris SA. (1995). Scaling ADLs within the MDS. J Gerontol A Biol Sci Med Sci. 1999;54(11):M546–M553.

Morris, J., Fries, B. E., Steel, K., Ikegami, N., Bernabei, R., & Carpenter, I., et al (1997). Comprehensive clinical assessment in community setting: Applicability of the MDS-HC. Journal of the American Geriatrics Society, 45 1017-1024.

Morris, J.N., Hawes, C., Fries, B.E., et al. (1990). Designing the National Resident Assessment Instrument for Nursing Facilities. The Gerontologist 30(3):293-307, June 1990.

Narayan, V. M., Narr, K. L., Kumari, V., Woods, R. P., Thompson, P. M., Toga, A. W., & Sharma, T. (2007). Regional Cortical Thinning in Subjects With Violent Antisocial Personality Disorder or Schizophrenia. The American Journal of Psychiatry, 164(9), 1418–1427. http://doi.org/10.1176/appi.ajp.2007.06101631

National Institute for Clinical Excellence (2004). Short-term Management of Violent (Disturbed) Behaviour in Adult Psychiatric In-patient and Accident and Emergency Settings. Second Draft for Consultation. London: NICE.

National Institute for Clinical Excellence (2015). National Collaborating Centre for Mental Health. Violence and aggression: short-term management in mental-health, health and community settings. NICE guideline [NG10].Update. Leicester and London, UK: British Psychological Society and The Royal College of Psychiatrists, 2015.

Needham, I., Abderhalden, C., Halfens, R. J., Fischer, J. E. & Dassen, T. (2005). Non-somatic effects of patient aggression on nurses: A systematic review. Journal of Advanced Nursing, 49, 283–296.

Neufeld, E., Perlman, C., & Hirdes, J. (2012). Predicting Inpatient Aggression Using the InterRAI Risk of Harm to Others Clinical Assessment Protocol : A Tool for Risk Assessment and Care Planning. The journal of behavioral health services & research (Vol. 39). https://doi.org/10.1007/s11414-011-9271-x

Ng B, Kumar S, Ranclaud M , et al. (2001). Ward crowding and incidents of violence on an acute inpatient unit. Psychiatric Services 52:521–525, 2001.

Nijman, H. L., Palmstierna, T., Almvik, R. et al. (2005b). Fifteen years of research with the Staff Observation Aggression Scale: A review. Acta Psychiatrica Scandinavica, 111, 12–21.

Nijman, H., Bowers, L., Oud, N. & Jansen, G. (2005a). Psychiatric nurses' experiences with inpatient aggression. Aggressive Behaviour, 31, 217–227.

OECD/EU (2013). A Good Life in Old Age?: Monitoring and Improving Quality in Long-term Care, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264194564-en. Accessed January 10th, 2018.

Otto R. (2000). Assessing and managing violence risk in outpatient settings. Journal of Clinical Psychology 56, 1239–1262.

Panksepp J. (1998). Affective Neuroscience: the Foundations of Human and Animal Emotion. New York, NY: Oxford University Press.

Perlman C, Barbaree H, Brown G, et al. Harm to others clinical assessment protocol (2011). In Hirdes JP, Curtin-Telegdi N, Mathias K, et al. interRAI Mental Health Clinical Assessment Protocols (CAPs) for Use with Community and Hospital-Based Mental Health Assessment Instruments. Version 9.1. Washington, DC: interRAI.

Perlman, C. M., & Hirdes, J. P. (2008). The aggressive behavior scale: A new scale to measure aggression based on the minimum data set. Journal of the American Geriatrics Society, 56(12), 2298-2303. 10.1111/j.1532-5415.2008.02048.x

Perlman, C. M., Hirdes, J. P., Barbaree, H., Fries, B. E., McKillop, I., Morris, J. N., & Rabinowitz, T. (2013). Development of mental health quality indicators (MHQIs) for inpatient psychiatry based on the interRAI mental health assessment. BMC Health Services Research, 13(1), 15. https://doi.org/10.1186/1472-6963-13-15

Quintal, S. A. (2002). Violence against psychiatric nurses: An untreated epidemic? Journal of Psychosocial Nursing and Mental Health Services, 40(1), 46-55.

Raine, A., Lencz, T., Bihrle, S., LaCasse, L., & Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. Archives of General Psychiatry, 57(2), 119–127. Retrieved from http://dx.doi.org/10.1001/archpsyc.57.2.119

Rajmohan, V., & Mohandas, E. (2007). The limbic system. Indian Journal of Psychiatry, 49(2), 132–139. http://doi.org/10.4103/0019-5545.33264

Rocca, P., Villari, V., & Bogetto, F. (2006). Managing the aggressive and violent patient in the psychiatric emergency. Progress in Neuropsychopharmacology & Biological Psychiatry, 30(4), 586-598. 10.1016/j.pnpbp.2006.01.015

Rosell, D., & Siever, L. (2015). The neurobiology of aggression and violence. CNS Spectrums, 20(3), 254-279. doi:10.1017/S109285291500019X

Royal College of Psychiatrists (1998). Management of Imminent Violence. Clinical Practice Guidelines to Support Mental Health Services (Occasional Paper OP41). London: Royal College of Psychiatrists

Sah P, Faber ES, Lopez De AM, Power J. (2003). The amygdaloid complex: anatomy and physiology. Physiol Rev.; 83(3): 803–834.

Salzman, C. D., & Fusi, S. (2010). Emotion, Cognition, and Mental State Representation in Amygdala and Prefrontal Cortex. Annual Review of Neuroscience, 33, 173–202. http://doi.org/10.1146/annurev.neuro.051508.135256

Serper M.R., Goldberg B.R., Herman K.G., et al. (2005). Predictors of aggression on the psychiatric inpatient service. Comprehensive Psychiatry 46, 121–127.

Shaffer, D. K., Kelly, B., & Lieberman, J. D. (2010). An Exemplar-Based Approach to Risk Assessment: Validating the Risk Management Systems Instrument. Criminal Justice Policy Review, 22(2), 167–186. https://doi.org/10.1177/0887403410372989

Siever, L. J. (2008). Neurobiology of Aggression and Violence. The American Journal of Psychiatry, 165(4), 429–442. http://doi.org/10.1176/appi.ajp.2008.07111774

Silver, E. (2000). Extending social disorganization theory: A multilevel approach to the study of violence among persons with mental illnesses. Criminology, 38(4), 1043–1074.

Soliman A.E.D. & Reza H. (2001) Risk factors and correlates of violence among acutely ill adult psychiatric inpatients. Psychiatric Services 52, 75–80.

Spencer S, and Johnson P. (2016). De-escalation techniques for managing aggression (Protocol). Cochrane Database of Systematic Reviews 2016, Issue 1. Art. No.: CD012034. DOI: 10.1002/14651858.CD012034.

Steadman, H. J., Silver, E., Monahan, J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., . . . Banks, S. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. Law and Human Behavior, 24(1), 83-100. 10.1023/A:1005478820425

Steinert T., Wölfle M. & Gebhardt R.-P. (2000). Measurement of violence during in-patient treatment and association with psychopathology. Acta Psychiatrica Scandinavica 102, 107–112.

Swanson, J. W., Holzer, C. E., Ganzu, V. K., et al. (1990). Violence and psychiatric disorder in the community: Evidence from the epidemiological catchment area surveys. Hospital and Community Psychiatry, 41, 761--770.

Tedeschi, J. T., & Felson, R. B. (1994). Violence, aggression, and coercive actions. Washington, DC: American Psychological Association.

Thimbleby, H. (2013). Technology and the Future of Healthcare. Journal of Public Health Research, 2(3), e28. http://doi.org/10.4081/jphr.2013.e28

Thomas, S., Leese, M., Walsh, E., McCrone, P., Moran, P., Burns, T., … Fahy, T. (2005). A comparison of statistical models in predicting violence in psychotic illness. Comprehensive Psychiatry, 46(4), 296–303. https://doi.org/https://doi.org/10.1016/j.comppsych.2004.10.001

Tomaino, J. J. (2010). "Understanding MDS 3.0 and RUG IV Reimbursement for Nursing Homes". Retrieved from https://rsmus.com/pdf/understanding_mds_rugiv_nursing_homes.pdf Retrieved Jan 12th 2018.

Tonkonogy JM, Geller JL. (1992). Hypothalamic lesions and intermittent explosive disorder. The Journal of Neuropsychiatry and Clinical Neurosciences, 4(1), 45–50. https://doi.org/10.1176/jnp.4.1.45

Tucker, A.M. (2009). Resource Utilization Groups (RUGs), Maryland LTC-PAC Workgroup, September 2, 2009. Retrieved Feb 6th from http://www.hilltopinstitute.org/publications/ResourceUtilizationGroups-RUGs-September2009.pdf

U.S. Department of Health and Human Services, Health Care Financing Administration, Long Term Care Resident Assessment Instrument User's Manual Version 2.0, October, 1995. Retrieved Jan 12th 2018

U.S. Department of Health and Human Services, Office of Inspector General (OIG), Nursing Home Resident Assessment Resource Utilization Groups, January, 2001. Retrieved Jan 12th 2018

Urbanoski, K. A., Mulsant, B. H., Willett, P., Ehtesham, S., & Rush, B. (2012). Real-World Evaluation of the Resident Assessment Instrument-Mental Health Assessment System. The Canadian Journal of Psychiatry, 57(11), 687–695. https://doi.org/10.1177/070674371205701106

Warburton, W., & Anderson, C. (2015). Aggression, Social Psychology of. International Encyclopedia of the Social & Behavioral Sciences. https://doi.org/10.1016/B978-0-08-097086-8.24002-6

Waters, Everett & Corcoran, David & Anafarta, Meltem. (2005). Attachment, Other Relationships, and the Theory that All Good Things Go Together. Human Development - HUM DEVELOP. 48. 80-84. 10.1159/000083217.

Webster, C. K., Douglas, D. E., Eaves, D., & Hart, D. (1997). HCR-20 assessing risk for violence: Version II. Burnaby, British Columbia, Canada: Mental Health, Law & Policy Institute, Simon Fraser University.

Wellens NI, Deschodt M, Flamaing J, Moons P, Boonen S, Boman X, Gosset C, Petermans J, Milisen K (2011). First-generation versus third-generation comprehensive geriatric assessment instruments in the acute hospital setting: a comparison of the minimum geriatric screening tools (MGST) and the interRAI acute care (interRAI AC). J Nutr Health Ageing, 15: 638-644. 10.1007/s12603-011-0109-2.

Whitley, G.; Jacobson, G.; Gawrys, M. (1996). The Impact of Violence in the Health Care Setting upon Nursing Education. Journal of Nursing Education, 35(5), 211–218

Whittington, R., & Wykes, T. (1996). "Going in strong": Confrontive coping by staff. Journal of Forensic Psychiatry 5, 609–614

Wong, S. C. P., & Gordon, A. (2006). The validity and reliability of the Violence Risk Scale: A treatment-friendly violence risk assessment tool. Psychology, Public Policy, and Law, 12, 279–309. doi:10.1037/1076-8971.12.3.279

World Health Organization [WHO] (2002). World Report on Violence and health: Summary. Geneva, WHO; 2002

Yamauchi, K. (1997). "Designing a New Payment System for Psychiatric Care: Development of a Case-Mix Classification System." Journal of the Japanese Society on Hospital Administration. 34(2): 155-200.

Yang, M., Wong, S., & Coid, J. (2010). The Efficacy of Violence Prediction: A Meta-Analytic Comparison of Nine Risk Assessment Tools. Psychological bulletin (Vol. 136). https://doi.org/10.1037/a0020473

Yesavage, J.A., T.L. Brink, T.L. Rose et al. (1982). "Development and Validation of a Geriatric Depression Screening Scale: A Preliminary Report", Journal of Psychiatric Research 198283, Vol. 17, No. 1, pp. 37-49.

Zillmann D. (1988). Cognition-excitation interdependencies in aggressive behavior. Aggress. Behav. 14:51–64

Zimmerman, D.R., S.L. Karon, G. Arling, B.R. Clark, T. Collins, R. Ross and F. Sainfort (1995). "Development and Testing of Nursing Home Quality Indicators", Health Care Financing Review, Vol. 16, No. 4, pp. 107-127, Summer.