Effects of Implementing the Short-Term Assessment of Risk and Treatability for Mechanical Restraint in a Forensic Male Population: A Stepped-Wedge, Cluster-Randomized Design

Jacob Hvidhjelm[1]*, Mette Brandt-Christensen[2], Christian Delcomyn[2], Jette Møllerhøj[3], Volkert Siersma[4] and Jesper Bak[1]

[1] Clinical Mental Health and Nursing Research Unit, Mental Health Center Sct Hans, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark, [2] Mental Health Centre Sct Hans, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark, [3] Head of Centre, Competence Centre for Forensic Psychiatry, Mental Health Centre Sct Hans, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark, [4] The Research Unit for General Practice and Section of General Practice, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

The assessment and formulation of the risk of violence and other unwanted behaviors at forensic psychiatric facilities have been attempted for decades. Structured professional judgment tools, such as the Short-Term Assessment of Risk and Treatability (START), are among the recent attempts to overcome the challenge of accomplishing these goals. This study examined the effect of implementing START in clinical practice for the most serious adverse events among the target group of severely mentally ill forensic psychiatric inpatients. Results were based on the use of mechanical restraints as an outcome. This study is a pragmatic, stepped-wedge, cluster-randomized controlled trial and was conducted over 5 years. It included eight forensic psychiatric units. Fifty out of 156 patients who had a basic aggression score of more than 0 were included in the study. We found that the rate of mechanical restraint use within the START period were 82% [relative risk (RR) = 0.18], lower than those outside of the START period. Patients evaluated within the START period were also found to have a 36% (RR = 0.64) lower risk of having higher Brøset Violence Checklist scores than patients evaluated outside the START period. Previous studies on START have primarily focused on validation, the predictive capability of the assessment, and implementation. We were only able to identify one study that aimed to identify the benefits and outcomes of START in a forensic setting. This study showed a significant reduction in the chance for inpatients in a forensic psychiatric facility to become mechanically restrained during periods where the START was used as risk assessment.

Keywords: mental health, psychiatry, forensic, coercion, mechanical restraint, risk assessment, Short-Term Assessment of Risk and Treatability

INTRODUCTION

Valid and reliable measures to assess the risk of violence and other challenging behaviors at forensic psychiatric facilities have been in demand for decades, and several structured professional judgment tools have been developed and introduced into clinical practice (1–8). The Short-Term Assessment of Risk and Treatability (START), a 20-item structured professional judgment instrument designed for recurrent clinical assessments within inpatient and community contexts, is among the most recently developed assessment tools (9). The predictive validity of START for several problem behaviors is generally considered good to excellent within a short to moderate timeframe (10–14). The implementation of START in clinical practice to focus on the patient's strengths and vulnerabilities has been assumed to provide enhanced opportunities to predict and prevent severe violence and self-harm. Scientific reports of using START alone or comparing START with other risk assessment tools in secure mental health settings have been published (11, 15–18). Previous research has primarily focused on its validation and predictive ability (19, 20). The most recent study focused on summarizing item values as a single concept, primarily for research purposes (21). To our knowledge, no previous study has highlighted the reduction in mechanical restraint use as an outcome. Since 2010, there has been an increased focus among Danish politicians and health authorities on reducing the use of coercion in hospital psychiatric departments. In 2014, the Danish government ordered a reduction of mechanical restraint use in inpatient settings by half before 2020, signaling that this was considered the most intrusive type of coercion applied. Additionally, it was a political goal that all types of coercion should decrease during this period. Considerable efforts and resources have been applied, leading to a reduction in the use of coercion; however, the overall goal of reducing the use of mechanical restraints by 50% compared to baseline (years 2011–2013) was not fully achieved by the end of 2020 (22). The present study was initiated by implementing START in clinical practice between May 1, 2012, and April 30, 2017, at a large, medium-secure, forensic mental health facility in the Capital Region of Denmark. All forensic psychiatric facilities in Denmark are publicly funded and are subject to public health authorities. This study aimed to examine the effects of START implementation in clinical practice on the most serious adverse events as expressed by the necessity for mechanical restraint use among the target group of severely mentally ill forensic psychiatric inpatients.

MATERIALS AND METHODS

Study Design

Randomized controlled trials are considered the gold standard when evaluating the effectiveness of interventions in the healthcare context. However, randomized controlled trials have some weaknesses in a workplace setting. Therefore, to overcome these challenges, we decided to apply a stepped-wedge, cluster-randomized design, which has some advantages over classic randomized controlled trials in this setting (23, 24). We attempted to overcome some of the difficulties associated with

the use of the stepped-wedge design. During our study, it was possible to gradually implement the intervention in all participating units and motivate the staff and patients.

Another gold standard is following the intention-to-treat principle. This is not preferable in a naturalistic scenario, such as ours, because it is not always possible to rescreen the patients within the maximum effect period of START (which we determined to be 6 months). Furthermore, because some units lost key staff members, they could not evaluate patients with START until new key staff members were trained. Therefore, if we had followed the intention-to-treat principle, we would not have been able to evaluate the effect of START; instead, we would have obtained the effect of the ability of the unit to perform START.

This study used the definition of mechanical restraint as defined by Bowers et al. (25) (the use of restraining straps, belts, or other equipment to restrict movement). This definition refers only to the restraint of inpatients at psychiatric hospitals. The following conditions must be present to legally initiate mechanical restraint according to the Danish Mental Health Act (as translated by the authors):

"Mechanical restraint may be used only when necessary to prevent patients from the following: (1) Exposing their body or health or the body or health of others to danger. (2) Pursuing or in any other way grossly molesting fellow patients. (3) Committing significant acts of vandalism" (26).

According to the law in Denmark, all coercive episodes must be reported to the national database for coercion (26).

Other Instruments
Staff Observation Aggression Scale-Revised

All aggressive or violent incidents were systematically recorded using the electronic version of the Staff Observation Aggression Scale-Revised (SOAS-R). The SOAS-R is an instrument that reports damaging or threatening aggressive behaviors toward an object and/or humans. The SOAS-R is completed each time a staff member witnesses aggressive or violent behavior by a patient. The SOAS-R has been tested and validated by several studies (27–29). With the SOAS-R scoring system, the severity of an incidence can be rated from 0 to 22 points; a score >8 is considered severe. The SOAS-R has shown good inter-rater reliability, with kappa values of 0.61–0.74 (27). The SOAS-R has been used in daily clinical practice since 2008. Staff is trained to register a SOAS-R whenever they witness or are themselves exposed to a violent incident. All registrations are entered in an IT system (designed by Frenzs B.V., Nijmegen, the Netherlands).

Brøset Violence Checklist

The Brøset Violence Checklist (BVC) is used to evaluate the presence (score of 1) or absence (score of 0) of six symptoms: confusion, irritability, boisterousness, physical threats, verbal threats, and attacks on objects. According to standard guidelines (30), a total score of 0 (none of these behaviors present) suggests that the risk of violence is low. A score of 1–2 suggests that the risk is moderate and preventive measures should be taken. A score of 3 or more suggests that the risk of violence is high, immediate preventive measures are required, and plans for

managing an attack should be activated (30, 31). The BVC was implemented in 2005 and recorded in patients files on daily basis.

Population and Timeframe

This study was conducted over 5 years, from May 1, 2012, to April 30, 2017, at the Forensic Department of the Mental Health Centre Sct Hans, Mental Health Services, in the Capital Region of Denmark. Ten units comprise the Forensic Department; however, one unit was excluded because it had served as a pilot unit, and one was excluded because it did not use mechanical restraints. The mean number of beds per unit was 9.4 (range, 8–10 beds).

We included all male forensic patients who displayed one or more basic aggressive episodes. A basic aggressive episode is defined as an episode involving a total SOAS-R score of more than 8 during the first month of inclusion in the study. A total of 50 male patients were included. The reason for excluding patients without one or more basic aggressive episodes was associated with applying mechanical restraints. In Denmark, mechanical restraints are only initiated if the patients are aggressive (toward themselves, others, or things). Therefore, if implementing the START would reduce the use of mechanical restraints, then it would only be possible to detect if the patients had aggression issues. To select patients with aggression issues, we selected those who experienced one or more severe aggressive episodes during the first month of inclusion in the study based on the assumption that those would be the ones most at risk for requiring mechanical restraints.

All units admitted both male and female patients. A total of 13 female patients were admitted during the study period (only three with a basic aggressive episode score were in the included eight units). However, they were excluded because they were presumed to have different associations between START and mechanical restraint use compared to the male patients and because they comprised a sample too small for separate analysis.

Sampling and Data Collection

Of the 10 units of the Forensic Department, eight were used for the study. Five units were randomized to step one: beginning the training for key staff to teach them how to screen patients using START on May 22, 2013. Two units were randomized to step two: beginning the training for key staff 1 year later, on May 21, 2014. The last unit began the training for key staff 2.5 years after step two had begun, on September 9, 2016. All data were retrospectively gathered in May 2017 (see **Figure 1**). The included units were randomized by one of the researchers (JB), using the random number generator in the statistics software that was used.

Key staff members (nurses) were trained in leading the START assessment meeting with multidisiplanery staff attending (e.g., nurses, assistant nurses, psychologists, and psychiatrist). The initial START assessment is often more time consuming than the follow-up assessments and therefore there was a natural decrease in the time spent doing a START assessment—from 1.5 h down to 30 min.

The vast majority of patients admitted to a forensic unit is admitted under court order. In Denmark, forensic psychiatry is part of general psychiatry and not a specialty in itself. The overall responsibility for initiating and implementing treatment is placed upon the treating psychiatrist and always happens in collaboration with other staff. A total of 239 patients were admitted during the study period between May 1, 2012, and April 30, 2017. After the first process of excluding patients from both the pilot unit, and the unit that did not use coercive measures, we were left with 169 patients who were assessed for eligibility. Based on the argumentation above, 13 female patients were additionally excluded. The remaining 156 patients were then rated based on their Basic Aggressive Episodes (Basic Aggressive Episode "BAE": episode involving a SOAS-R score >8 during the first month of participation in the study). A total of 50 (36%) patients who displayed one or more BAE were included in the study. The START period began when the patients underwent the first START evaluation, and it proceeded until the patients did not undergo the START evaluation for 6 months. After that time, they began the control period; however, another START period could have begun if another START evaluation had been performed. This procedure resulted in 42 START periods and 92 control periods for these patients (see **Figure 2**).

Outcome Variables

The primary outcome was the occurrence of mechanical restraint use. Secondary outcomes were the total duration (in minutes) of mechanical restraint use, total coercive episodes (number of physical restraint episodes, episodes of acute forced medication, episodes of one to one observation (without patient consent) and mechanical restraint episodes) and the number of BVC scores more than 2 (a score of 3–6 indicated a severe risk of violence within the next 24 h). The BVC scores were determined three times every 24 h. In a Danish context, mechanical restraint is considered the most severe type of coercion, as the use of seclusion rooms is not allowed. Therefore, mechanical restraint was selected as the outcome measure.
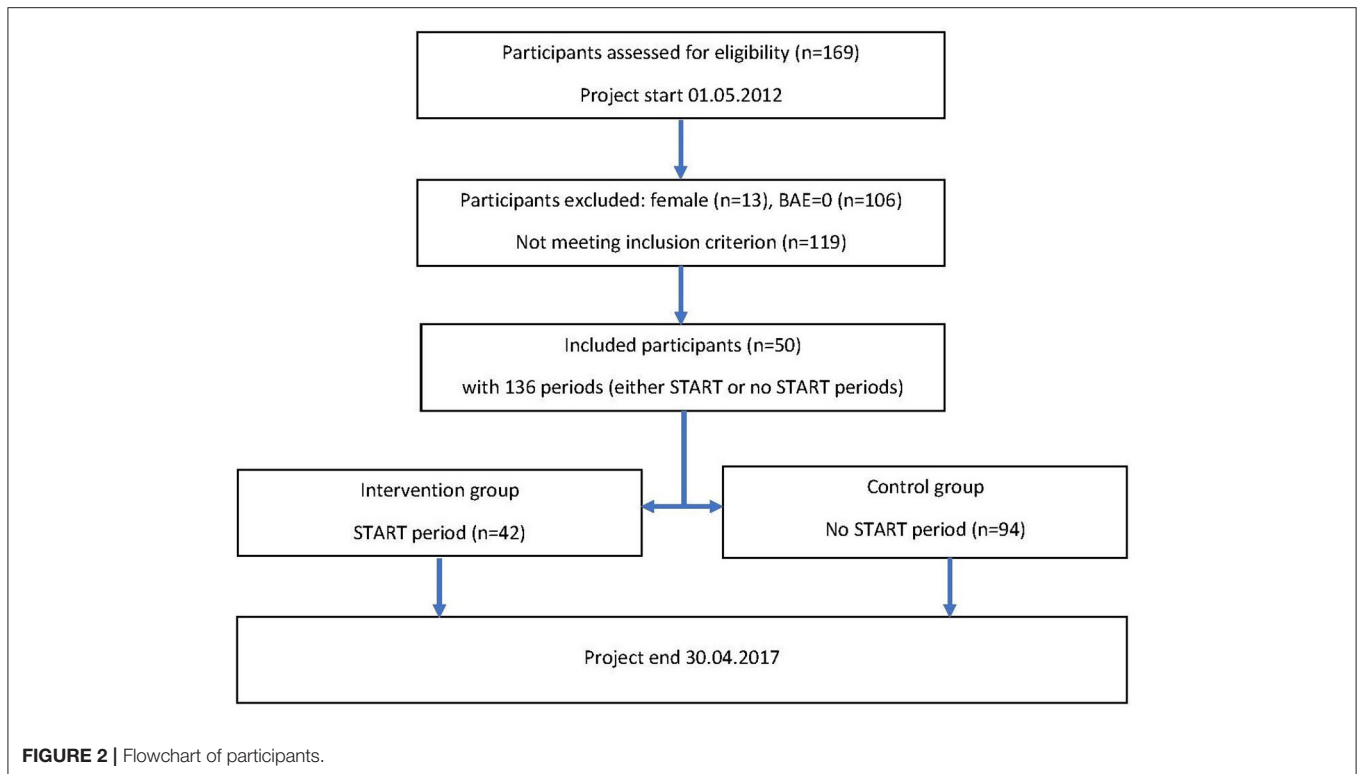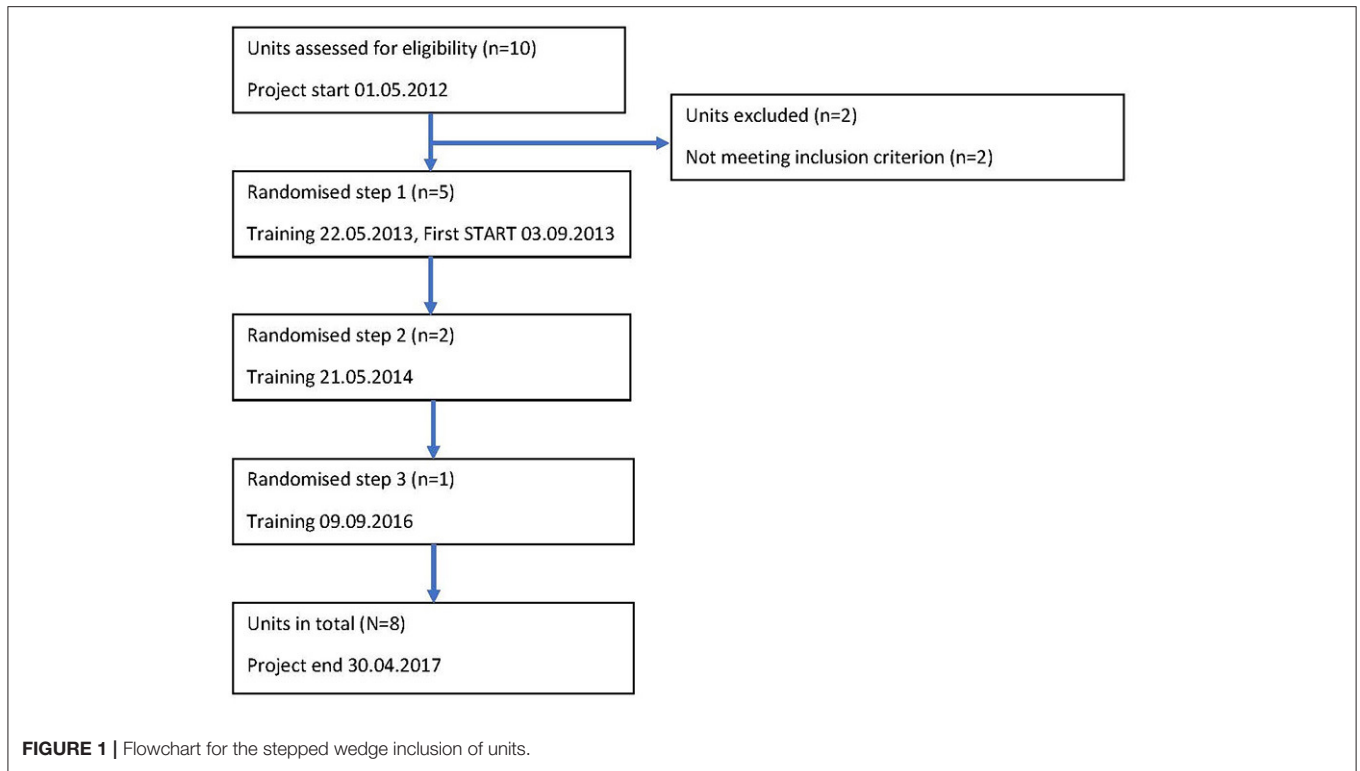
We selected mechanical restraint *(coercion)* rather than SOAS-R as our primary outcome measure. The SOAS-R outcome variable is known to have a relative high degree of underreporting (32). As mentioned earlier registration is mandatory by law and therefore we assume much less underreporting on mechanical restraint than on the SOAS-R.

Potential Confounders

Normally, randomization eliminates the effect of potential confounding variables caused by even distribution. To ensure this in our study, we gathered information about the most important potential confounding variables for this group of patients: age, diagnosis, length of hospitalization before study inclusion, and psychoactive substance use.

Previous Study Findings

According to previous studies, younger men required mechanical restraint use more often, and patients diagnosed with schizophrenia or schizotypal and delusional disorders (WHO ICD-10-codes F20-F29) required mechanical restraint use more often (33). Patients required mechanical restraint use more often at the beginning of their hospitalization period (33, 34). Further,

FIGURE 1 | Flowchart for the stepped wedge inclusion of units.

FIGURE 2 | Flowchart of participants.

patients diagnosed with mental and behavioral disorders because of psychoactive substance (WHO ICD-10-codes F10-F19) use experienced mechanical restraint use more often (33).

Data Analysis

Negative binomial regression was performed to assess the incidence of mechanical restraint use when START evaluations

were performed compared to when no START evaluations were performed (35). To assess a cluster effect of "unit", we calculated the proportion of the explained variance attributable to a "unit" as the difference of the coefficients of determination (R^2) of a model with and without "unit" included as categorical variable, divided by the R^2 of a model with "unit" included. The most important covariates were tested for differences in the BAE groups, the START group, and the control group using the chi-squared test. However, because there were both paired and unpaired data, the Cochran-Mantel-Haenszel test was also performed. We analyzed the patient-level data to eliminate the cluster effect. Statistical significance was set at $p < 0.05$. All analyses were performed using SPSS Statistics for Windows version 25.0 (36).

Approval and Ethics

Interventions such as those included in the present study do not require approval from the scientific ethical committee system in Denmark, because it does not include any drugs or biological material and the intervention is regarded part of the natural improvement of care and treatment. The Danish Data Protection Agency (RHP-2013-002, I-Suite no. 02053) approved this study. We received permission from the Center Management and Clinical Management of the Forensic Department to perform this study. We also received permission from the developers of START to use their method during this study (April 1, 2014). This study followed the ethical guidelines for nursing research in the Nordic countries (37) and the recommendations on the legal protection of persons suffering from mental disorders, especially those placed as involuntary patients (38).

RESULTS

Patients who were 28 to 35 years of age had a higher prevalence of having one or more basic aggressive episodes at the beginning of the study period (BAE > 0, 22.0% vs. BAE = 0, 32.1%), but the difference was not significant ($p = 0.20$). Additionally, patients with one or more basic aggressive episodes at the beginning of the study period (basic aggressive episodes >0) had a higher prevalence of schizophrenia, schizotypal or delusional disorder, and other delusional disorders (F20-F29) (BAE > 0, 84.0% vs. BAE = 0, 75.5%), but the difference was not significant ($p = 0.23$) (**Table 1**).

A total of 296 (72.7%) START assessments (on both patients with a BAE score = 0 and BAE above 0) out of 407 was preformed during the study period with a mean of 4.1 month between assessments. One patient had 17 START assessments preformed during the study period.

Patients in the intervention group (START period) had been hospitalized before being included in the study ($p = 0.01$) (**Table 2**). The Cochran-Mantel-Haenszel test results were similar for these patients.

The rate of mechanical restraint use within the START period was 82% lower than that outside the START period [relative risk (RR) = 0.18; 95% confidence interval (CI), 0.08–0.41; $p < 0.01$; $p = 3.01 \times 10^{-5}$]. The results of an adjusted analysis (RR = 0.17; 95% CI, 0.08–0.37; $p < 0.01$; $p = 6.0 \times 10^{-6}$) were similar, indicating that incidence differences could not be explained by confounding factors. The proportion of fit attributable to "unit" was 0.04 (4%) (**Table 3**).

The duration of mechanical restraint use was 99% lower within the START period than outside the START period (RR = 0.01; 95% CI, 0.00–0.01; $p < 0.01$; $p = 2.0 \times 10^{-14}$). The results of an adjusted analysis (RR = 0.002; 95% CI, 0.001–0.006; $p < 0.01$; $p = 0.0 \times 10\text{-E}$) were similar, indicating that the incidence difference could not be explained by confounding factors. The proportion of fit attributable to "unit" was 0.01 (1%). The very small RR, could probably be explained by a few patients, mechanical restrained for a long period of time. Therefor, the analyses of duration should not be the primary result (**Table 3**).

The rate of total use of coercion within the START period was 63% lower than that outside the START period [relative risk (RR) = 0.37; 95% confidence interval (CI), 0.19–0.74; $p < 0.01$]. The results of an adjusted analysis (RR = 0.33; 95% CI, 0.19–0.60; $p < 0.00$) were similar, indicating that incidence differences

TABLE 1 | Descriptive statistics of background variables of the whole population.

	BAE = 0 (n = 106)		BAE > 0 (n = 50)		Total (N = 156)		χ^2
	n	%	*n*	%	*N*	%	*p*-value
Age when the patient was included in the study							
<27 years	28	26.4%	16	32.0%	44	28.2%	0.47
28–35 years	34	32.1%	11	22.0%	45	28.8%	0.20
36–45 years	25	23.6%	13	26.0%	38	24.4%	0.74
>45 years	19	17.9%	10	20.0%	29	18.6%	0.76
Diagnoses of schizophrenia, schizotypal and delusional disorders (F20–F29)	80	75.5%	42	84.0%	122	78.2%	0.23
Length of hospitalization before the patient were included in the study (<1 year)	55	51.9%	29	58.0%	84	53.8%	0.48
Diagnoses of psychoactive substance use, primary or secondary (F10–19)	68	64.2%	32	64.0%	100	64.1%	0.99

BAE (Basic Aggressive Episodes) = the sum of SOAS-R scorings above 8, the first month the patient were in the study. Patients with at least one BAE (BAE > 0) was included in the study. χ^2, Chi square test for the difference (Δ) between the BAE = 0 group and the BAE > 0 group.

TABLE 2 | Descriptive statistics of the background variables of the START group and the control group.

	START group (n = 42)		Control group (n = 94)		Total (N = 136)		χ^2
	n	%	n	%	n	%	p-value
Age when the patient was included in the study							
<27 years	11	26.2%	26	27.7%	37	27.2%	0.86
28–35 years	9	21.4%	19	20.2%	28	20.6%	0.87
36–45 years	12	28.6%	28	29.8%	40	29.4%	0.89
>45 years	10	23.8%	21	22.3%	31	22.8%	0.85
Diagnoses of schizophrenia, schizotypal and delusional disorders (F20–F29)	39	92.9%	82	87.2%	121	89.0%	0.33
Length of hospitalization before the patient were included in the study (<1 year)	13	31.0%	51	54.3%	64	47.1%	0.01*
Diagnoses of psychoactive substance use, primary or secondary (F10–19)	25	59.5%	57	60.6%	82	60.3%	0.53

A START period is the day and time of the START rating plus 6 months. A control period is all other periods where the patient is hospitalized. The 50 patients (BAE > 0) had a total of 136 periods and all patients in the START group also was a part of the control group. 22 patients only participated in the control group with 29 periods. χ^2, Chi square test for the difference between the START group and the control group.
*$p \leq 0.05$.

TABLE 3 | Effect of START on the outcome variables: mechanical restraint use, total coercion and BVC episodes.

	START group (n = 42)		Control group (n = 94)		Unadjusted analysis[1] (N = 136)				Adjusted analysis[2] (N = 136)				Proportion of fit attributable to "unit"[3]
	Rate (#/month)		Rate (#/month)		RR	95% Wald confidence interval of RR		p-value	RR	95% Wald confidence interval of RR		p-value	
	Mean	SD	Mean	SD		Lower	Upper			Lower	Upper		
MR[4] episodes	0.03	0.10	0.18	0.39	0.18	0.08	0.41	0.00**	0.17	0.08	0.37	0.00**	0.04
Duration in minutes of MR[4] episodes	274	1652	609	1950	0.01	0.00	0.03	0.00**	0.00	0.00	0.01	0.00**	0.01
Total coercive episodes[5]	0.12	0.32	0.39	0.94	0.37	0.19	0.74	0.01*	0.33	0.19	0.60	0.00**	0.01
Number of BVC[6] episodes (>2)	.96	1.38	1.42	1.72	0.64	0.44	0.91	0.01*	0.60	0.43	0.86	0.01**	0.08

[1] Negative Binominal Regression, Offset = log. to the length of the period, Repeated Subject = Patient ID, Adjusted for Units (cluster effect).
[2] Further adjusted for: Age, Diagnoses, Length of hospitalization, and Psychoactive substance use.
[3] Proportion of fit attributable to "unit" = the proportion of R2 with and without units (clusters).
[4] Mechanical Restraint.
[5] Total coercive episodes = number of physical restraint episodes, episodes of acute forced medication, episodes of one to one observation (without patient consent) and mechanical restraint episodes.
[6] BVC, Brøset Violence Checklist.
*$p \leq 0.05$.
**$p \leq 0.01$.

could not be explained by confounding factors. These results (RR = 0.37 compared to RR = 0.18), indicates that some of the mechanical restraint episodes is converted to a lesser intrusive kind of coercion (which in it self would be a positive result), but not to a degree that impacts the results especially. The proportion of fit attributable to "unit" was 0.01 (1%) (**Table 3**).

Finally, the risk of having a BVC score more than 0 (BVC score >2 indicated a severe risk of violence within the next 24 h) within the START period was 36% lower than that outside the START period (RR = 0.64; 95% CI, 0.44–0.91; p = 0.01). The results of

an adjusted analysis (RR = 0.60; 95% CI, 0.43–0.86; p = 0.01) were similar, indicating that the incidence difference could not be explained by confounding. The proportion of fit attributable to "unit" was 0.08 (8%) (**Table 3**).

DISCUSSION

The stepped-wedge design was chosen for both ethical and practical reasons. Ethical considerations included the absence of equipoise because there is evidence that START will do more

good than harm. Therefore, it would be unethical to withhold its implementation from participants. Practically, the design solved the problem of simultaneous implementation in half the units and many logistical and practical problems (39, 40). Additionally, the outcomes were available from routinely collected data (BVC and SOAS-R) (40). The design also allowed us to include a large number of patients. The clustering of sites confined to one geographical site and the number of sites could have been limitations to the generalizability of the current study (35, 39, 41).

In an ideal world, where the time between START assessments had been possible to keep as recommended in the manual (9), and where key employees that was responsible for arranging START assessments meetings was not ill or had left their position in the department, and that sufficient resources were available so that the team performing the assessment had opportunity to meet. We would have expected to have collected 407 START assessments but collected 296 assessments. It might not even be possible to reach a complesion rate of 73% in this ideal world, as we were able to reach in our naturalistic scenario.

Previous research has primarily focused on the validation and predictive ability of START. Few studies have focused on its implementation and outcomes, especially in inpatient settings. Kroppan et al. (11) described START implementation in two phases at Forensic Mental Services at Brøset in Trondheim, Norway, during a 10-year period. Their study showed increased interdisciplinary participation with the implementation of START. The research group also highlighted that the implementation of START requires continuous efforts. The application of the assessments to the treatment plans proved challenging when the study was performed, although progression from the assessment to the assessment-treatment phase during the implementation period was identified. In the current study, the START implementation was performed by clinicians who had experience using START in the clinical setting and experience training clinicians to use START. In each unit, two superusers were educated about START and were in charge of its implementation in their unit. They were supervised by the superusers of the pilot unit, who were supervised by teachers from Norway. The superusers met during the implementation phase to rate cases together. The intensive use of external supervision and continuous follow-up could explain the positive implementation process and significant study results.

To our knowledge, the study by Gunenc et al. (17) is the only one that focused on the benefits and outcomes of START in a forensic inpatient setting. They expected to find a reduction in adverse behaviors (physical and verbal aggression, self-harm, victimization, self-neglect, unauthorized leave, and substance abuse); however, they found no significant changes in physical or verbal aggression over time. There was no reduction in self-harm or substance abuse incidents during the 3 months after the START evaluation. Despite the power calculations, the authors (17) indicated that the sample size ($n = 50$) was one explanation for their results. During our study, we opted for a study period that was considerably longer than the 3-month comparison period before and after the assessment follow-up period used by Gunenc et al. (17). The longer study period might be one of the reasons for our significant results. We only included patients who

demonstrated that they can use aggression to express themselves; therefore, patients were included if they had a BAE score >0.

According to the research literature within the field of forensic psychiatry and personal recovery processes, START is emphasized as an important risk assessment tool because it focuses on the resources, strengths, and protective factors in addition to the weaknesses and risks of the individual patients (42, 43). Managing risk as well as positive risk-taking and protective factors are key offender recovery elements in specialized forensic services, and this implies the involvement of mentally disordered offenders in their risk assessment and management and reduction of specific risks (44–46). Consequently, it would be relevant to develop a patient version of START to support and increase the involvement of the mentally disordered offenders, thereby supporting the processes of personal recovery. Lockertsen et al. (47) added items to the original version of the BVC and studied their extended version; for example, their Self-Report Risk Scale provided patients with an opportunity to predict their risk of violence. As a result, their study showed that expressing one's risk resulted in better short-term accuracy of predicting violence than the original BVC (45).

In a systematic review by Goulet et al. (48) its being concluded that with implementation of a program that focuses on reducing seclusion and reduction it is possible to affected the use of such methodes in a positive way. The review defines such programs as programs including the following key-components; Leadership, training, post-seclusion and restraint review, patient involvement, use of prevention tools and focus on the therapeutic environment. The use of START is one component in one of the mentioned six key-components, namely "Prevention tools". In our study the only implemented or used component that in daily practice separates the intervention group and the control group is the use of START.

A patient version of START would probably help to increase patient awareness of risk factors and highlight the responsibilities of the patients and the professionals working with them. As pointed out in the substantial literature about recovery processes in forensic settings, such approaches can help patients regain a sense of control over their lives, thereby providing hope (39). Furthermore, the implementation of START in specialized forensic outpatient services would be an interesting area for future research in the context of Denmark. Troquete et al. (49) examined the preventive effect of combining START and a shared care protocol in forensic outpatient settings without finding a significant preventive effect on recidivism to violent or criminal behavior. They (49) stated that the proportion of clients in the intervention group not receiving the intervention or receiving it only once was a limitation. Additionally, they did not have much success motivating the case managers to perform activities during their study, which was a limitation to their study (49). Therefore, this topic requires further scrutiny.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

JH, CD, and JB carried out the project. VS supervised the analysis. JH and JB took lead in writing the manuscript. All authors discussed the results and contributed to the final manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

REFERENCES

1. Chu CM, Thomas SDM, Ogloff JRP, Daffern M. Short—the short—to medium-term predictive accuracy of static and dynamic risk assessment measures in a secure forensic hospital. *Assessment.* (2013) 20:230–41. doi: 10.1177/1073191111418298

2. Crocker AG, Garcia A, Israel M, Hindle Y, Gagnon D, Venegas C, et al. *Implementing and using a systematic risk assessment scheme to increase patient safety on a risk management unit for individuals with severe mental illness: A demonstration project.* Edmonton, AB: Canadian Patient Safety Institute (2008).

3. Daffern M. The predictive validity and practical utility of structured schemes were used to assess the risk for aggression in psychiatric inpatient settings. *Aggression Violent Behav.* (2007) 12:116–30. doi: 10.1016/j.avb.2006.03.005

4. Dickens, GL, O'Shea LE, Christensen M. Structured assessments for imminent aggression in mental health and correctional settings: systematic review and meta-analysis. Int J Nurs Stud. (2020) 104–104:103526. doi: 10.1016/j.ijnurstu.2020.103526

5. Hogan NR, Olver ME. Assessing risk for aggression in forensic psychiatric inpatients: An examination of five measures. *Law Hum Behav.* (2016) 40:233–43. doi: 10.1037/lhb0000179

6. Lewis AHO, Webster CD. General instruments for risk assessment. *Curr Opin Psychiatry.* (2004) 17:401–5. doi: 10.1097/01.yco.0000139978.68060.5a

7. Singh JP, Serper M, Reinharth J, Fazel S. Structured assessment of violence risk in schizophrenia and other psychiatric disorders: a systematic review of the validity, reliability, and item content of 10 available instruments. *Schizophr Bull.* (2011) 37:899–912. doi: 10.1093/schbul/sbr093

8. Webster CD, Douglas KS, Eaves D, Hart S. *HCR-20: Assessing the Risk of Violence.* Vancouver: Simon Fraser University Mental Health Law and Policy Institute (1997).

9. Webster C, Martin M, Brink J, Nicholls T, Desmarais S. *Manual for the Short-termTerm Assessment of Risk and Treatability (START). 1.1 Version.* Coquitlam: British Columbia Mental Health and Addiction Services (2009).

10. Cartwright JK, Desmarais SL, Hazel J, Griffith T, Azizian A. Predictive validity of HCR-20, START, and static-99R assessments in predicting institutional aggression among sexual offenders. *Law Hum Behav.* (2018) 42:13–25. doi: 10.1037/lhb0000263

11. Kroppan E, Nonstad K, Iversen RB, Søndenaa E. Implementation of the short-term assessment of risk and treatability over two phases. *J Multidiscip Healthc.* (2017) 10:321–6. doi: 10.2147/JMDH.S133514

12. Paetsch A, Van Os TWDP, Troquete NAC, Van Den Brink RHS. Single-item predictive validity of the short-term assessment of risk and treatability (START) for violent behavior in outpatient forensic psychiatry. *J Forensic Psychiatry Psychol.* (2019) 30:630–41. doi: 10.1080/14789949.2019.1581831

13. Troquete NA, van den Brink RH, Beintema H, Mulder T, van Os TW, Schoevers RA, et al. Predictive validity of the short-term assessment of risk and treatability for violent behavior in outpatient forensic psychiatric patients. *Psychol Assess.* (2015) 27:377–91. doi: 10.1037/a0038270

14. Whittington R, Bjørngaard JH, Brown A, Nathan R, Noblett S, Quinn B. Dynamic relationship between multiple START assessments and violent incidents over time: a prospective cohort study. *BMC Psychiatry.* (2014) 14:323. doi: 10.1186/s12888-014-0323-7

15. Abidin Z, Davoren M, Naughton L, Gibbons O, Nulty A, Kennedy HG. Susceptibility (risk and protective) factors for in-patient violence and self-harm: prospective study of structured professional judgment instruments START and SAPROF, DUNDRUM-3, and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry.* (2013) 13:197. doi: 10.1186/1471-244X-13-197

16. Dickens GL, O'Shea LE. How short should a short-term risk assessment be? Determining the optimum interval for START reassessment in a secure mental health service. *J Psychiatr Ment Health Nurs.* (2015) 22:397–406. doi: 10.1111/jpm.12232

17. Gunenc C, O'Shea LE, Dickens GL. Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: use of the START. *Crim Behav Ment Health.* (2018) 28:61–71. doi: 10.1002/cbm.2036

18. O'Shea LE, Dickens GL. Predictive validity of the START for unauthorised leave and substance abuse in a secure mental health setting: a pseudo-prospective cohort study. *Int J Nurs Stud.* (2015) 52:970–9. doi: 10.1016/j.ijnurstu.2015.02.007

19. O'Shea LE, Dickens GL. Short-term assessment of risk and treatability (START): systematic review and meta-analysis. *Psychol Assess.* (2014) 26:990–1002. doi: 10.1037/a0036794

20. Ramesh T, Igoumenou A, Montes MV, Fazel S. Use of risk assessment instruments to predict violence in forensic psychiatric hospitals: a systematic review and meta-analysis. *European psychiatry.* (2018) 52:47–53. doi: 10.1016/j.eurpsy.2018.02.007

21. Whittington R, Pollak C, Keski-Valkama A, Brown A, Haines-Delmont A, Bak J, et al. Unidimensionality of the strengths and vulnerabilities scales in the short-term assessment of risk and treatability (START). *Int J Forensic Ment Health.* (2021) 20:1–10. doi: 10.1080/14999013.2021.1953193

22. Danish Health Authority. *Anbefalinger for nedbringelse af tvang for mennesker med psykiske lidelser.* København: Sundhedsstyrelsen (2021).

23. Hemming K, Taljaard M, Grimshaw J. Introducing the new CONSORT extension for stepped-wedge cluster randomized trials. *Trials.* (2019) 20:68. doi: 10.1186/s13063-018-3116-3

24. Brown CA, Lilford RJ. The Stepped wedge trial design: a systematic review. *BMC Med Res Methodol.* (2006) 6:54. doi: 10.1186/1471-2288-6-54

25. Bowers L, Van Der Werf B, Vokkolainen A, Muir-Cochrane E, Allan T, Alexander J. International variation in containment measures for disturbed psychiatric inpatients: a comparative questionnaire survey. *Int J Nurs Stud.* (2007) 44:357–64. doi: 10.1016/j.ijnurstu.2006.01.005

26. Bekendtgørelse af lov om anvendelse af tvang i psykiatrien m.v. *Psykiatriloven Den Sundheds Ældreministeriet LBK* 936 af 02/09/2019. Available online at: https://www.retsinformation.dk/eli/lta/2019/936 (accessed February 9, 2022).

27. Nijman HL, Palmstierna T, Almvik R, Stolker JJ. Fifteen years of research with the staff observation aggression scale: a review. *Acta Psychiatr Scand.* (2005) 111:12–21. doi: 10.1111/j.1600-0447.2004.00417.x

28. Nijman HLI, Muris P, Merckelbach HLGJ, Palmstierna T, Wistedt B, Vos AM, et al. The staff observation aggression scale-revised (SOAS-R). *Aggr Behav.* (1999) 25:197–209. doi: 10.1002/(SICI)1098-2337(1999)25:3<197::AID-AB4>3.0.CO;2-C

29. Palmstierna T, Wistedt B. Staff observation aggression scale, SOAS: Presentation and evaluation. *Acta Psychiatr Scand.* (1987) 76:657–63. doi: 10.1111/j.1600-0447.1987.tb02936.x

30. Almvik R, Woods P. The brøset violence checklist (BVC) and prediction of inpatient violence: preliminary results. *Psychiatr Care.* (1998) 5:208–13.

31. Almvik R, Woods P, Rasmussen K. The brøset violence checklist: sensitivity, specificity, and interrater reliability. *J Interpers Violence.* (2000) 15:1284–96. doi: 10.1177/088626000015012003

32. Hvidhjelm J, Sestoft D, Bjørner JB. The Aggression observation short form identified episodes not reported on the staff observation aggression scale-revised. *Issues Ment Health Nurs.* (2014) 35:464–9. doi: 10.3109/01612840.2013.879359

33. Thomsen C, Starkopf L, Hastrup LH, Andersen PK, Nordentoft M, Benros ME. Risk factors of coercion among psychiatric inpatients: A nationwide register-based cohort study. *Soc Psychiatry Psychiatr Epidemiol.* (2017) 52:979–87. doi: 10.1007/s00127-017-1363-3

34. Andersen K, Nielsen B. Coercion in psychiatry: the importance of extramural factors. *Nord J Psychiatry.* (2016) 70:606–10. doi: 10.1080/08039488.2016.1190401

35. Hooper R, Copas A. Stepped wedge trials with continuous recruitment require new ways of thinking. *J Clin Epidemiol.* (2019) 116:161–6. doi: 10.1016/j.jclinepi.2019.05.037

36. Corportation I, IBM. *SPSS Statistics for Windows (version 25.0, Armonk, NY, USA).* Armonk, NY: IBM Corp (2017).

37. Northern Nurse F. Ethical guidelines for nursing research in the Nordic Countries. *Vard Nord.* (2003) 23:1–19.

38. De L'Europe C. *"White Paper" on the Protection of the Human Rights and Dignity of People Suffering From Mental Disorder, Especially Those Placed as Involuntary Patients in a Psychiatric Establishment.* Srasbourg: Council of Europe (2000).

39. Zhan Z, Van Den Heuvel ER, Doornbos PM, Burger H, Verberne CJ, Wiggers T, et al. Strengths and weaknesses of a stepped wedge cluster randomized design: Its application in a colorectal cancer follow-up study. *J Clin Epidemiol.* (2014) 67:454–61. doi: 10.1016/j.jclinepi.2013.10.018 (accessed February 9, 2022).

40. Beard E, Lewis JJ, Copas A, Davey C, Osrin D, Baio G, et al. Stepped wedge randomised controlled trials: Systematic review of studies published between 2010 and 2014. *Trials.* (2015) 16:353. doi: 10.1186/s13063-015-0839-2

41. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ.* (2015) 350:h391. doi: 10.1136/bmj.h391

42. Drennan G, Wooldridge J, Aiyegbusi A, Alred D, Ayres J, Barker R, et al. *Making Recovery a Reality in Forensic Settings, ImROC Briefing Paper 10.* London: Centre for Mental Health Mental Health Network/NHS Confederation (2014).

43. Simpson AIF, Penney SR. Recovery and forensic care: recent advances and future directions. *Crim Behav Ment Health.* (2018) 28:383–389. doi: 10.1002/cbm.2090

44. Hillbrand M, Young JL, Griffith EE. Managing risk and recovery: redefining the miscibility of oil and water. *J Am Acad Psychiatry Law.* (2010) 38:452–456. doi: 10.1002/clen.201000044

45. Drennan G, Alred D. Secure recovery: approaches to recovery in forensic mental healthmental health settings. *Routledge.* (2012). 23–40. doi: 10.4324/9780203129173

46. Møllerhøj J. Offender recovery. forensic patient perspectives on long-term personal recovery processes. *Int J Environ Res Public Health.* (2021) 18:6260. doi: 10.3390/ijerph18126260

47. Lockertsen Ø, Varvin S, Færden A, Eriksen BMS, Roaldset JO, Procter NG, et al. Risk assessment of imminent violence in acute psychiatry: a step towards an extended model. *J Forensic Psychiatry Psychol.* (2020) 31:41–63. doi: 10.1080/14789949.2019.1663898

48. Goulet M-H, Larue C, Dumais A. Evaluation of seclusion and restraint reduction programs in mental health: a systematic review. *Aggress Violent Behav.* (2017) 34:139–46. doi: 10.1016/j.avb.2017.01.019

49. Troquete NA, Van Den Brink RH, Beintema H, Mulder T, van Os TW, Schoevers RA, Wiersma D. Risk assessment and shared care planning in out-patient forensic psychiatry: Cluster randomized controlled trial. *Br J Psychiatry.* (2013) 202:365–371. doi: 10.1192/bjp.bp.112.113043