Implementation of guidelines on prevention of coercion and violence (PreVCo) in psychiatry: a multicentre randomised controlled trial

Tilman Steinert,[a,b,*] Johanna Baumgardt,[c] Andreas Bechdolf,[c] Felix Bühling-Schindowski,[c] Celline Cole,[d] Erich Flammer,[a,b] Susanne Jaeger,[a,b] Julia Junghanss,[d] Marie Kampmann,[a,b] Lieselotte Mahler,[d] Rainer Muche,[e] Dorothea Sauter,[a] Angelika Vandamme,[d] and Sophie Hirsch[a,b]

[a]Ulm University, Germany
[b]Centres for Psychiatry Suedwuerttemberg, Germany
[c]Vivantes Klinikum am Urban, Germany
[d]Charité Berlin University, Germany
[e]Ulm University, Institute of Epidemiology and Medical Biometry, Germany

The Lancet Regional Health - Europe 2023;35: 100770

Published Online xxx https://doi.org/10.1016/j.lanepe.2023.100770

Summary

Background Interventions to prevent the use of coercion in psychiatric hospitals have been summarized in the 2018 German Association for Psychiatry, Psychotherapy, and Psychosomatic's comprehensive guidelines. Twelve recommendations for implementation of these guideline on psychiatric wards have been deducted and their feasibility has been tested in a pilot study, using external implementation consultants as facilitators. The objective of the PreVCo study was to test their effect in a randomised clinical trial.

Methods Fifty-four psychiatric wards in Germany treating voluntary and involuntary patients were randomly allocated to either an intervention or to a waiting list condition. The intervention consisted of the implementation of three out of 12 suggested recommendations as selected by the ward teams, supported by external study workers. As the primary outcome measure, the number of coercive measures used per bed and month in the final 3 months of the intervention period was determined. Secondary outcomes were the cumulative duration of coercive measures used per bed and months and assaults per bed and month. Achieved guideline adherence was measured by a fidelity scale developed for this purpose during a pilot study for the PreVCo Rating Tool. After a 3-month baseline collection period under routine conditions, randomisation was done after matching wards pairwise according to frequency of coercive measures used and scores on the PreVCo Rating Tool at baseline. The duration of the intervention period was 12 months; control wards received only an initial workshop presentation of the study and completed their PreVCo ratings. We used the Wilcoxon signed rank test and the paired t-test and conducted sensitivity analyses for different periods of observation.

Findings Neither the number of coercive measures used per month and bed nor their cumulative duration nor the number of assaults per bed and months differed significantly between the 27 intervention wards and the 27 control wards in the final 3 months of the intervention period. The median number of coercive measures used decreased by 45% (median 0.96 (IQR 1.34)–0.53 (IQR 0.59) from baseline until the end of the intervention period on the intervention wards and by 28% (median 0.98 (IQR 1.71)–0.71 (IQR 1.08) on waiting list wards. The PreVCo Rating Tool showed a significant improvement in intervention wards compared to control wards, indicating a successful implementation.

Interpretation The study demonstrated that guideline adherence could be significantly improved by the intervention. However, there was no evidence for an effect on the frequency or duration of coercive measures used. Spill-over effects and the impact of the COVID-19 pandemic on in-patient care might have limited the effect of the intervention. Further research from robust randomised controlled trials are necessary to identify effective interventions to reduce the use of coercion in psychiatric hospitals.

Funding The study was funded by the German *Innovationsfonds beim Gemeinsamen Bundesausschuss* (project no. 01VSF19037). The funder had no role in study design or data collection.

DOI of original article: https://doi.org/10.1016/j.lanepe.2023.100795

*Corresponding author. Klinik für Psychiatrie und Psychotherapie I der Universität Ulm (Weissenau), Weingartshofer Str. 2, D 88214, Ravensburg, Germany.
E-mail address: tilman.steinert@zfp-zentrum.de (T. Steinert).

Keywords: Coercion; Seclusion; Restraint; RCT; Violence; Implementation

Research in context

Evidence before this study

To detect high quality, multisite RCTs on interventions to reduce the use of coercion in mental health services we conducted a PubMed search with the term ([coerc*] AND [multi*]) AND (randomized) (18.10.2023) and included available reviews on interventions to prevent or reduce the use of coercion. We found mixed evidence and a considerable number of studies showing positive effects, but they were exclusively observational studies (e.g., pre-post design) or clearly underpowered RCTs without a power calculation. The small number of multicentre RCTs with sufficient statistical power all yielded negative results, for example on the effect of joint crisis plans, involuntary outpatient commitment, or involvement of peer workers. Also, many strategies and programmes that have been implemented more or less widely in clinical practice are not yet supported by robust evidence on patient-related outcomes so far. While a large number of observational studies with pre-post-design report a substantial reduction of the use of coercion, negative results have been reported for outpatient coercive treatment settings in the United Kingdom, joint crisis plans in the UK and Germany, a guideline-based multicomponent intervention in nursing homes in Germany, and standardized post-coercion debriefings in Germany. Some cluster-randomized RCTs that were clearly underpowered reported positive results for risk assessment, Six Core Strategies, and de-escalation. We could identify only one sufficiently powered RCT with a positive result in terms of reduction of involuntary readmissions—a recent study from France on advance directives facilitated by peer workers. Interestingly, also for the internationally most well-known complex interventions that are part of our 12 recommendations—Six Core Strategies and Safewards—no well-designed RCT proving efficacy of the use of coercion has been published so far, notwithstanding a large number of studies with lower evidence levels and mixed evidence. Furthermore, we searched for studies and reviews on implementation of guidelines in psychiatric services by the search term (implement*[Title]) AND (guideline [Title]) AND (psychiatr*) AND ([random*] OR [review]) (18.10.2023). Research on the effect of guidelines is still rather scarce and evidence is mixed. According to a systematic review, effects were mostly small, absent, or even deterioration was observed in terms of provider performance, but somewhat better in terms of patient-related outcomes.

Added value of this study

We could not provide evidence that implementation of guideline-based recommendations reduces the use of coercive measures in psychiatric hospitals. However, the findings indicate that the approach of using external experts as facilitators might be an effective strategy to implement complex interventions in multi-professional psychiatric teams.

Implications of all the available evidence

Preventing coercion in psychiatric hospitals is considered very important in terms of both quality of treatment and human rights. A wide variety of interventions and programmes have been evaluated and/or introduced in clinical practice in many countries and are summarized in guideline recommendations. However, robust evidence of their efficacy is still lacking. Further evidence from well-designed, sufficiently powered RCTs is necessary.

Introduction

Nowadays, there is a general consensus that coercive measures such as restraint or seclusion of a patient should only be used as a last resort if, despite de-escalation attempts and offers of treatment, there is still an acute danger to the patient or others.[1] In Germany, the Federal Constitutional Court decided in 2018 that each intervention/measure using a mechanical restraint lasting longer than 30 min requires an assessment by a judge at the patient's bedside. This decision led to an adaptation of national and federal mental health laws, and the percentage of patients subjected to any kind of freedom-restricting coercion decreased from 6.6% in 2017 to 5.8% in 2019, while the median cumulated duration of restraint and seclusion per affected case decreased from 12.5 to 11.9 h.[2] Considering the power of the Court's mandate, these effects were highly significant but moderate. In line with this finding, observational data from 2004 to 2019 showed that, in spite of many efforts to reduce coercion and awareness programmes, the observed reduction in the percentage of affected cases was predominantly due to changed practices in old age psychiatry, while in general psychiatry there was not much change.[3]

Therefore, from the view of the German Society for Psychiatry, Psychotherapy, and Psychosomatics (DGPPN), stronger and more clearly defined interventions were necessary to reduce the use of coercion. In 2018, the DGPPN's clinical practice guidelines *Prevention of Coercion: Prevention and Treatment of Aggressive Behaviour in Adults*[4] were published. The guidelines were developed in a well-defined, formal way

with comprehensive research evidence as well as a structured consensus process with 22 societies and stakeholders, including those of service users and their relatives. Eventually, the guidelines comprise 89 recommendations and statements based on evidence and achieved consensus, referring to prevention of coercion and violence and treatment of aggressive behaviour. Part of the guidelines was a systematic review on measures to avoid coercion in psychiatry.[5] It had turned out that there is limited evidence for several interventions to be conducted on psychiatric wards. However, for none of them there is evidence from randomized controlled multicentre trials with sufficient power. Thus, current guideline recommendations are based on a mix of (limited) evidence and strong consensus. The guidelines are available as a comprehensive book[4] and also as clinical practice guidelines that were sent to all psychiatric hospitals by the DGPPN, and they are also available free of charge in full-text on the Internet.

However, publishing guidelines with recommendations does not yet mean they will be followed in practice. In the field of research on coercion, there is increasing consensus that the next important step will not be to develop completely new interventions but to implement those that have proven effective in real-world practice.[6] Both the steep costs involved in developing high-quality guidelines and the low benefits if application in practice is realized insufficiently have been the subject of repeated criticism. Concerns have even been voiced by guideline authors themselves in recent years.[7] Generally, there is a relative paucity of research on implementation in psychiatry and available results are not very encouraging.[8] While studies have not shown a consistent positive effect of guideline implementation on provider performance, a more consistent small to modest positive effect on patient outcomes has been observed.[8]

Thus, we identified two research gaps. First, there is a lack of well-designed controlled studies with sufficient power to test the efficacy of interventions to reduce the use of coercion on psychiatric wards. Second, there is a lack of evidence as to how complex psychiatric guidelines can be successfully implemented in psychiatric practice. Consequently, we conducted a research programme in several steps, of which the first was the development of the guidelines. Next, we derived 12 recommendations for implementation on the ward level, six additional ones for implementation on the hospital level, and eight additional recommendations for implementation on the community level.[9] These recommendations were approved by the DGPPN in 2018. We focused on guideline implementation on psychiatric wards via the 12 recommendations and decided for a strategy to use consultants who were well-known experts in the field. We tested the feasibility of this implementation strategy in a pilot study with six psychiatric wards.[9] For the pilot study, we developed a fidelity scale as an outcome measure for provider performance, the PreVCo Rating Tool. This scale and the counselling approach used in the present study were subsequently adapted according to the results of the pilot study. The main study that is reported here was subsequently designed to test the efficacy of the intervention.

The primary study question was whether the number of coercive measures used per bed and month in psychiatric wards could be reduced by implementation of the recommendations of the PreVCo programme. A secondary objective of the study was to evaluate the chosen approach of external implementation consultants in terms of improving adherence to the recommendations.

Methods

Study design

A detailed description of the study design and methods has been published in the study protocol.[10] This was a mixed methods study design. The comprehensive results of the qualitative research on facilitators and barriers for the implementation will be reported elsewhere. The trial was registered in the Clinical Trial Registration (www.isrctn.com) with the identifier ISRCTN 71467851. It is reported according to SPIRIT guidelines.[11] Sample size was calculated for the primary outcome assuming an effect size of 0.6, a power of >80% and a significance level of 5% (for details see[10]). In this multicentre study, 54 wards were randomized in a 1:1 ratio to either an intervention or a control condition (waiting list), stratified by the amount of coercive measures used per bed and month and implemented aspects of the guidelines to matched pairs. In addition, a waiting-list control design allows for a pre-post analysis for all participating wards. Furthermore, this design allows for analysing if observer effects already lead to a reduction of coercive measures between the baseline and the start of the intervention as well as for assessing potential spill-over effects in the control group during the waiting time. After 12 months, control wards received the intervention. The sample size calculation is described in detail in the methods supplement.

Ethical approval

The primary ethical approval was obtained from the Ethical Committee of the Ulm University, and, subsequently, from the responsible ethical committees for the participating hospitals.

Inclusion criteria

Wards were eligible for the study if they were responsible for admissions of involuntary patients, independent of diagnoses; if they collected data on coercive measures and aggressive incidents routinely; and if the medical and nursing management declared their support for the aims of the study. In the study protocol, we

had expressed the intention to include 52 psychiatric wards.[10] We recruited three additional wards, as we had concerns about losing participating wards due to the strains of the COVID-19 pandemic. Unexpectedly, this was not the case. Due to the odd number of participating wards, we excluded the ward that fitted least for matching.

Implementation concept

The intervention and its theoretical background are described in detail in the supplement according to the TIDieR Checklist.[12] The implementation was realised by means of trained implementation consultants as facilitators, incorporating insights from the preceding pilot study.[9] The aim of the intervention was to implement three out of a set of 12 recommendations or at least one complex intervention programme drawn from the guidelines, eventually selected by the ward teams according to their preferences, needs, and opportunities. Thus, this was a flexible, tailored approach. This procedure was chosen as there is evidence that tailored approaches are more effective than no intervention or dissemination of guidelines.[13] Implementation consultants held close contact to identified key persons (mostly nursing ward managers) and conducted workshops with the ward teams (T0 at baseline, T1 at the beginning of the intervention, T2 during the intervention, T3 after the intervention). Implementation and guideline adherence were assessed by the PreVCo Rating Tool[9] at baseline and at the end of the intervention period. The trial profile is displayed in Fig. 1.

Outcomes

The primary outcome was the number of coercive measures (seclusion, restraint, forced medication) used per occupied bed and month. Secondary outcomes were the cumulative duration of seclusion or restraint per occupied bed an month, and the change of the PreVCo Rating Tool after the intervention. In our previously published study protocol[10] we had planned to use data from all outcome measures as collected during the 12-month intervention period. However, due to circumstances resulting from the COVID-19 pandemic, some wards had different functions for part of the observation period or could not operate under their regular conditions for some time. Thus, valid data were available for some wards of both groups for less than 12 months. Therefore we made a study protocol amendment (see methods supplement) and explained that we used the data collected for the final 3 months of all wards for the primary analysis. These data were available completely for all wards. This approach had several advantages that could partly compensate for the flaws of the impact of the pandemic: (i) The length of the observation period was identical for all wards, and (ii) it can be assumed that the implementation of recommendations takes some time and will exert its impact mostly within the last 3 months of the intervention period. Accordingly, the study protocol was amended and the agreement of the funding agency was obtained. To compensate for the shorter period of analysis, we employed an additional sensitivity analysis using the complete time period with available valid data from each ward.

Data collection

The data collection is described in detail in the methods supplement.

Assessment of fidelity

To assess the degree of pre-existing adherence to the guidelines as well as change over time after the intervention period, we used the PreVCo Rating Tool.[9] In this instrument, each of the 12 recommendations is rated from 0 ("not implemented at all") to 9 ("fully implemented"). If a complex intervention like the Weddinger model,[14] Safewards,[15] or the Six Core Strategies[16] was chosen, the rating was multiplied by 4. Thus, a score on the PreVCo Rating Tool has a possible range between 0 and 135 points. PreVCo ratings were performed by the implementation consultants together with the respective ward teams in a consensus process during study visits at the beginning of the baseline period and after the end of the 12 month intervention or, respectively, after the 12 month waiting time period. The PreVCo Rating Tool had been developed and tested in a pilot study[9] and showed a moderate intra-class correlation (0.607; 95%-CI 0.364; 0.834). Further details are provided in the methods supplement.

Randomisation

The participating wards were matched in pairs following the best-fit principle according to the two baseline criteria *frequency of coercive measures per bed and month* and *initial score of the PreVCo Rating Tool*. To avoid spill-over effects, wards which belonged to the same hospital were allocated to the same randomisation arm. Thus, of the 55 participating wards, 54 wards were matched into 27 pairs. Afterwards, the wards were randomised into either intervention or waiting-list control groups by a 1:1 block randomisation. For details see[17] and methods supplement.

Statistical analyses

To assess for possible between-group differences in pairs, we employed Wilcoxon signed-rank tests, using a significance level of 0.05. The comparison of the outcomes of the final 3 months of the intervention period between both groups for the primary outcome was confirmatory, whereas all other comparisons and analyses were explorative. As a sensitivity analysis, we compared the complete 12-month period with available data from each ward using a per-protocol approach. All outcomes, except for PreVCo Rating Tool scores, were averaged over the final 3 months of the intervention

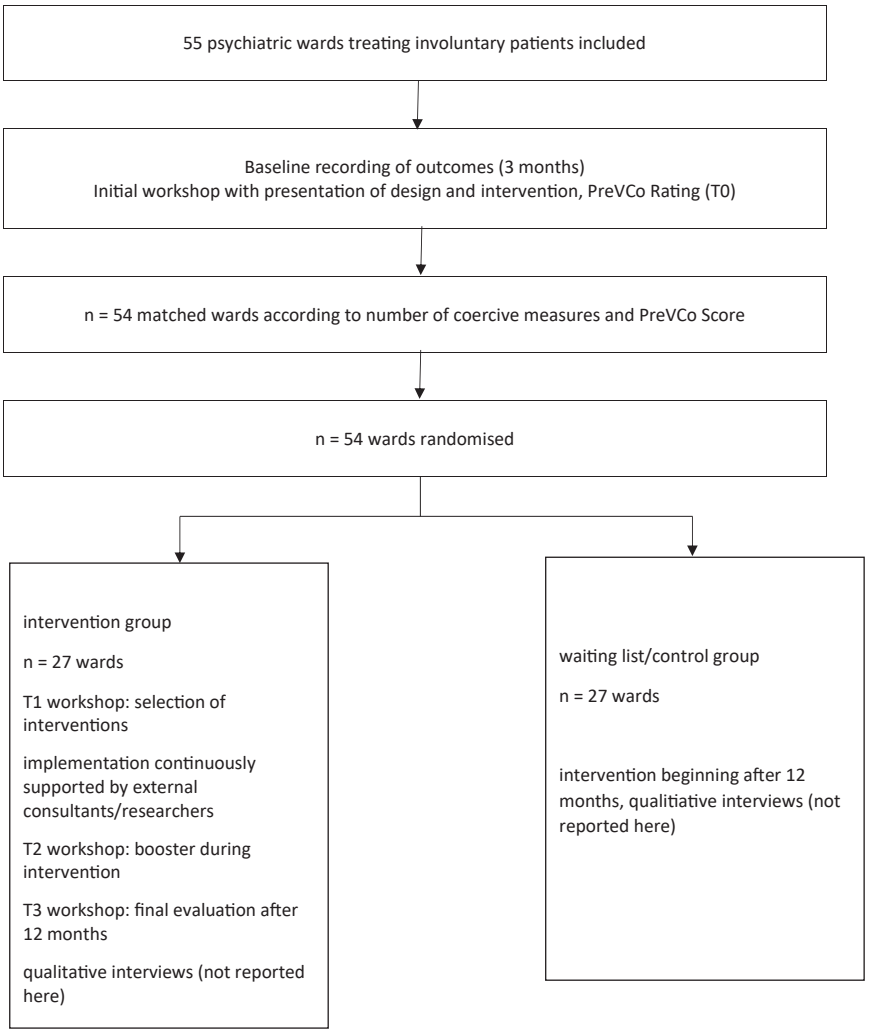

Fig. 1: **Trial profile**.

period. For the comparison of selected interventions with not-selected interventions, each PreVCo Rating Tool score was standardized by dividing the corresponding number of interventions. Due to the skewed distribution of data, secondary outcomes were also evaluated by the Wilcoxon signed-rank test and respective methods. The distribution of the PreVCo Rating Tool scores did not differ statistically significantly from a normal distribution and was therefore evaluated by paired t-tests. To assess the possible influence of covariates, we used general estimating equations (GEE) for the estimation of population-average effects. We used negative-binomial models with log link for the number of coercive measures per month and occupied and inverse normal models with log link. To protect against misspecification of the covariance matrix, we used robust sandwich estimators and to control for possible baseline differences, we used baseline measurements as covariate. Additionally, we performed a pre-post-analysis for the primary and secondary outcomes.

Data collection was done with Microsoft Excel 2013; data management was done with QlikTech QlikView 12; and statistical calculations were done with IBM SPSS 27. Effect size calculation was done with G*Power 3.1.9.6.

Role of the funding source

The study was funded by the German *Innovationsfonds beim Gemeinsamen Bundesausschuss* (project no. 01VSF19037). The funder had no role in study design or data collection.

Results

Baseline characteristics of the randomised wards

Baseline characteristics as recorded in the baseline period from June 1, 2020 until August 31, 2020 have

been published in detail previously[17] and are summarised in Table 1.

No significant differences were observed in the outcomes that were used for matching—the number of coercive measures used and the PreVCo Rating Tool scores. Only the number of recorded assaults per month was lower in the later intervention group.

Description of the randomised wards during the intervention period

In total, 28,118 patients were discharged during the 12-month intervention period. Characteristics of the wards as recorded during the 12-month intervention period are displayed in Table 2. Additional information is provided in the results supplement.

Selected interventions

Interventions selected by the wards after the initial workshop were monthly team meetings reviewing situations with coercive measures (41x), debriefing with patients after coercive measures (40x), architecture and environment (16x), risk prediction and management (13x), joint crisis plans (11x), complex interventions (11x), de-escalation training (9x), reliable recording of measures (8x), continuous 1:1 supervision during coercive measures, guideline-adhering pharmacotherapy, and implementation of peer work (each 2x). Establishing ward manuals on the use of coercive measures was not selected as an intervention.

Primary and secondary outcomes

After 12 months of intervention, there were no statistically significant differences between the intervention group and the waiting list in the number of coercive measures used per month and bed in the last 3 months (Table 3), the cumulated duration of coercive measures used per month and bed, and the number of assaults per month and bed. The primary outcome (i.e., the number of coercive measures used) was somewhat lower in the intervention group, but it failed to reach significance. A 45% decrease in the number of coercive measures used per bed and month was observed over time from the baseline to the intervention period for the intervention group (median 0.96 (IQR 1.34)–0.53 (IQR 0.59), $p = 0.046$). Also in the waiting list group an 28% decrease was observed (median 0.98 (IQR 1.71)–0.71 (IQR 1.08), $p = 0.016$). Multiple regression models confirmed that the primary outcome was significantly predicted only by the proportion of involuntary cases per month, but not by the condition (see results supplement).

No significant pre-post changes were observed in the cumulated duration of coercive measures used (intervention group $p = 0.43$, waiting list group $p = 0.79$) and the number of assaults (intervention group $p = 0.23$, waiting list group $p = 0.11$). The intervention group and the waiting list group differed statistically significantly on PreVCo Rating Tool scores ($p < 0.001$, Table 3), with a large effect size (Cohen's $d = 1.48$). PreVCo Rating Tool scores increased significantly over time in the

	Baseline data of waiting list wards $n = 27$	Baseline data of intervention wards $n = 27$	Differences between matched pairs $n = 27$	p (Wilcoxon signed- rank test)
Number of coercive measures used per month and occupied bed (primary outcome) Median (IQR)	0.98 (1.71)	0.96 (1.34)	−0.22 (0.74)	0.44[a]
Cumulated duration of coercive measures used per month and occupied bed (hours) Median (IQR)	6.61 (29.23)	7.22 (12.97)	0.03 (14.24)	0.61[a]
Number of assaults per month and occupied bed Median (IQR)	0.34 (0.57)	0.23 (0.57)	−0.06 (0.46)	0.049[a]
PreVCo Rating Tool score Mean (SD)	66.7 (16.8)	63.9 (16.9)	−2.8 (11.2)	0.202[b]

[a]Wilcoxon signed-rank test. [b]Paired t-test.

Table 1: **Baseline characteristics of primary and secondary outcomes.**

	Waiting list wards $n = 27$ Median (IQR)	Intervention wards $n = 27$ Median (IQR)
Occupied beds per month	19 (8)	18 (6)
Admissions per month	38 (19)	44 (18)
Percentage of involuntarily admitted cases per month	19 (31)	22 (21)
Number of nurses per month	17.1 (3.7)	15.5 (3.1)
Number of nurses per month and occupied bed	0.9 (0.3)	0.9 (0.4)
Number of doctors/psychologists per month	3.3 (1.1)	3.4 (1.2)
Number of doctors/psychologists per month and occupied bed	0.2 (0.1)	0.2 (0.1)

Table 2: **Characteristics of the randomised wards during the 12-month intervention period.**

	Waiting list wards n = 27	Intervention wards n = 27	Differences between matched pairs (intervention wards vs. waiting list wards) n = 27	p
Number of coercive measures used per month and occupied bed Median (IQR)	0.71 (1.08)	0.53 (0.59)	−0.043 (0.85)	0.68[a]
Cumulated duration of coercive measures used per month and occupied bed (hours) Median (IQR)	3.92 (25.00)	6.36 (16.94)	2.81 (31.30)	0.67[a]
Number of assaults per month and occupied bed Median (IQR)	0.32 (0.42)	0.19 (0.34)	−0.12 (0.52)	0.14[a]
PreVCo Rating Tool score Mean (SD)	66.2 (15.6)	78.6 (14.5)	12.4 (14.2)	0.00011[b]

[a]Wilcoxon signed-rank test. [b]Paired t-test.

Table 3: **Primary and secondary outcomes.**

intervention group ($p < 0.0001$), but not in the waiting list group ($p = 0.89$). In the multiple regression models, the condition was the only significant predictor of the PreVCo Rating Tool Score (see results supplement). Contrasting the selected interventions with those not selected in the intervention group, there was a significant increase from baseline to post intervention in standardized PreVCo Rating Tool scores for the selected interventions (median = 3.3 (IQR 1.8)–6.5 (IQR 2.0), $p < 0.0001$) and in the non-selected (median = 5.3 (IQR 1.5)–5.7 (IQR 1.2), $p = 0.00049$). The distribution of the data is displayed in Figs. S1–S12 in the results supplement.

The sensitivity analysis for primary and secondary outcomes (see results supplement) provided similar results.

Discussion

The major result of the PreVCo study was that we could not show that the intervention aiming at improving guideline adherence had a significant impact on the primary outcome measure, the frequency of coercive measures used in the participating wards. Compared to the baseline period, coercive measures decreased by 45% on the intervention wards, but also on the waiting list wards these measures decreased by 28%, while the number of assaults remained unchanged. Also in secondary outcome measures, cumulative duration of coercive interventions, and frequency of assaults, no significant differences emerged between intervention and control wards. This finding adds to a number of negative results from well-designed multicentre randomised controlled trials (RCTs) worldwide intended to reduce the amount of coercion in psychiatric services[18–22] with only one exception,[23] while RCTs with positive results were clearly underpowered or had other significant problems in methods.[24–27] In contrast, descriptive and pre-post studies yield considerably more positive results, as in many areas of health research.[28] Reducing coercion seems to be difficult to accomplish and perhaps rather easier to accomplish by legal changes than through clinical interventions.[5]

We identified several factors that might be responsible for the negative result of the study. First, unexpectedly, a considerable reduction of coercive measures occurred also on the waiting list wards. This was partly inevitable due to the study design. After study enrolment, an initial workshop was necessary to assess PreVCo Rating Tool scores at baseline. Most ward teams hoped to be randomised to the intervention wards because they were keen to start using the interventions they were presented with. After being randomised to the waiting list, some teams expressed their wish to start with some changes in their practice of the use of coercion. From the ethical perspective, wards could not be interdicted to make attempts to reduce the use of coercion due to study purposes. Second, the COVID-19 pandemic made it necessary to conduct a considerable part of the workshops in video-conference formats. This condition may have weakened personal commitment and team cohesion. Generally speaking, the focus of attention probably shifted from the reduction of coercion to the prevention of COVID-19 infections. Also, the number of admissions decreased, however, involuntary admissions and restrictive practices concomitantly increased,[29] as observed in many parts of Germany[30] and probably in most countries around the world. In addition, some wards could no longer complete their regular tasks for some time because they had to be transformed into specialized infection or quarantine wards. That said, the advantage of the RCT design is that these changes in clinical practice should theoretically have affected intervention wards and control wards similarly and simultaneously. It might be possible that participating ward teams invested less effort in avoiding coercion than they might have otherwise under non-COVID-19 conditions.

Third, the statistical power with pairs (n = 27) of wards was insufficient to detect small effects as observed here. The statistical power necessarily refers to the number of clinical wards, not to the much higher number of 27,726 treated patients. Fourth, the potential of the intervention could not yet be fully realised, as only three (or two, in the case of complex interventions) out of 12 suggested interventions had to be selected. Even

the three chosen recommendations might not have been fully implemented during the intervention period due to the adverse conditions of the pandemic. Also the character of tailored interventions implicated that not all wards might have received the same amount of interventions. Last, the average adherence to recommendations was already rather high at baseline with 64 (38–106) of a possible 135 points scored on the PreVCo Rating Tool, indicating a high awareness of the issue of coercion in psychiatric wards in Germany in recent years. Twenty years before, nearly all participating wards would probably have achieved a rating just above 0. Thus, ceiling effects might have played a role. On the other hand, a bottom effect also played a role since some wards recorded very low rates of coercion used at baseline, setting limits to all further improvements.

More encouraging were our results with respect to implementation, showing that it clearly worked, as measured by the PreVCo Rating Tool. Team performance in terms of guideline implementation was significantly better in the intervention group after 1 year. This finding contrasts results of a previous review that, surprisingly, guideline implementation had frequently either none, only weak, or even negative effects on guideline adherence.[8] We interpret our significant result from a 15-point PreVCo improvement with an effect size of d equalling 1.48 for the PreVCo Rating Tool in the intervention group as strong evidence for the success of our implementation strategy. Our novel strategy was built on specialist implementation consultants, using insights from our pilot study.[9] The clear improvement of PreVCo Rating Tool scores on the intervention wards was particularly remarkable in the light of the strains of the COVID-19 pandemic and the recurrent necessity to conduct workshops as video conferences.

Limitations

The limitations caused by the study design and the circumstances of the COVID-19 pandemic that might have contributed to the negative result were described above. Information on patient level was not recorded and it is not known which impact individual patient characteristics had on the results of the study. The evidence for the successful implementation of the recommendations is based on the consensual assessment of the ward teams and the external reviewer (the implementation consultant) with the PreVCo Rating Tool; however, there is no independent rating or objective data proving that adherence to guidelines really improved in everyday clinical practice. Anyway, according to the obtained ratings, the intervention was probably strong enough to put the selected recommendations into practice. The more general question is whether the recommendations by themselves can be effective. They comprise all available knowledge on prevention of coercion as detailed in guidelines, but it has to be conceded that many guideline recommendations are based on strong consensus rather than on robust evidence (e.g., joint crisis plans, peer involvement, debriefing, and environmental factors). At the end of the day, the provided guidelines represent the best tools we have presently and workers and ward teams felt that they had a strong role to play in improving the quality of clinical practice, respecting patients' dignity, and helping to maintain team coherence.

Conclusion

Further research from adequately powered, robust RCTs is necessary to identify effective interventions to reduce the use of coercion in psychiatric hospitals. The approach of using experienced external implementation consultants could be a successful strategy for guideline implementation, generally.

Contributors

TS framed the study design, the ethics application, and the funding application. Also, TS wrote the manuscript. SH and EF collected data from the wards and did the statistical analysis. SH supervised the entire study in terms of compliance with the study protocol. RM provided the statistical design, randomization, and statistical advice in the analysis of results. DS, CC, FB, and JJ performed the workshops at the wards, collected data from the wards, and employed the PreVCo Rating Tool. SJ and MK provided important input to the study design and interpretation of results. AB, JB, AV, and LM conducted the study in the northern and eastern parts of Germany, acquired and accompanied participating hospitals, and participated in the phrasing of manuscripts and interpretation of results. Members of the PreVCo Study Group are listed in the methods supplement.

Data sharing statement

The data are accessible under https://doi.org/10.5061/dryad.tht76hf5d.

Declaration of interests

TS declared that he had received funding for other research projects from government bodies and for the development of the guideline from the German Association for Psychiatry and Psychotherapy (DGPPN). All other authors declared that they had no conflicts of interests.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanepe.2023.100770.

References

1 European Committee for the Prevention of Torture and Inhuman or Degrading Treatment or Punishment (CPT). Means of Restraint in psychiatric Establishments for Adults 21/03/2017. https://rm.coe.int/16807001c3. Accessed December 15, 2022.
2 Flammer E, Hirsch S, Steinert T. Effect of the introduction of immediate judge's decisions in 2018 on the use of coercive measures in psychiatric hospitals in Germany: a population-based study. *Lancet Reg Health Eur*. 2021;11:100233.
3 Steinert T, Hirsch S, Goebel R, Snellgrove B, Flammer E. Reduction of coercive measures under routine conditions in psychiatric hospitals 2004-2019: strong effects in old age psychiatry, much less in general psychiatry. *Eur Psychiatry*. 2020;63:e102.
4 German Association for Psychiatry, Psychotherapy and Psychosomatics (DGPPN). *Clinical practice Guidelines on Prevention of Coercion in Adults' Psychiatry in Germany*. Berlin: Springer; 2019.
5 Hirsch S, Steinert T. Measures to avoid coercion in psychiatry and their efficacy. *Dtsch Arztebl Int*. 2019;116:336–343.
6 Barbui C, Purgato M, Abdulmalik J, et al. Efficacy of interventions to reduce coercive treatment in mental health services: umbrella review of randomised evidence. *Br J Psychiatry*. 2020;27:1–11.

7 Hoch E, Petersen K, Kopp I, Batra A, Mann K. Methodology of the S3 guidelines on alcohol and tobacco-related disorders [German]. *Nervenarzt*. 2016;87:6–12.
8 Girlanda F, Fiedler I, Becker T, Barbui C, Koesters M. The evidence-practice gap in specialist mental healthcare: systematic review and meta-analysis of guideline implementation studies. *Br J Psychiatry*. 2017;210:24–30.
9 Bechdolf A, Bühling-Schindowski F, Weinmann S, et al. DGPPN pilot study on the implementation of the S3 guideline "Prevention of coercion: prevention and therapy of aggressive behavior in adults"[German]. *Nervenarzt*. 2022;93:450–458.
10 Steinert T, Bechdolf A, Mahler L, et al. Implementation of guidelines on prevention of coercion and violence (PreVCo) in psychiatry: study protocol of a randomized controlled trial (RCT). *Front Psychiatry*. 2020;11:579176.
11 Chan A-W, Tetzlaff J, Altman D, et al. SPIRIT: new guidance for content of clinical trial protocols. *Lancet*. 2013;381:91–92.
12 Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. 2014;348:g1687.
13 Baker R, Camosso-Stefinovic J, Gillies C, et al. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev*. 2010;3:CD005470.
14 Mahler L, Heinz A, Jarchov-Jàdi I, Bermpohl F, Montag C, Wullschleger A. Therapeutic attitudes and structures in (open) acute psychiatry: the Wedding model [German]. *Nervenarzt*. 2019;90:700–704.
15 Baumgardt J, Jäckel D, Helber-Böhlen H, et al. Preventing and reducing coercive measures – an evaluation of the implementation of the Safewards model in two locked wards in Germany. *Front Psychiatry*. 2019;10:340.
16 Wieman DA, Camacho-Gonsalves T, Huckshorn KA, Leff S. Multisite study of an evidence-based practice to reduce seclusion and restraint in psychiatric inpatient facilities. *Psychiatr Serv*. 2014;65:345–351.
17 Hirsch S, Baumgardt J, Bechdolf A, et al. Implementation of guidelines on prevention of coercion and violence: baseline data of the randomized controlled PreVCo study. *Front Psychiatry*. 2023;14:1130727.
18 Burns T, Rugkåsa J, Molodynski A, et al. Community treatment orders for patients with psychosis (OCTET): a randomised controlled trial. *Lancet*. 2013;381:1627–1633.
19 Thornicroft G, Farrelly S, Szmukler G, et al. Clinical outcomes of joint crisis plans to reduce compulsory treatment for people with psychosis: a randomised controlled trial. *Lancet*. 2013;381:1634–1641.
20 Rixe J, Neumann E, Möller J, et al. Joint crisis plans and crisis cards in inpatient psychiatric treatment–a multicenter randomized controlled trial. *Dtsch Arztebl Int*. 2023;120(8):125–132. https://doi.org/10.3238/arztebl.m2022.0384.
21 Abraham J, Kupfer R, Behncke A, et al. Implementation of a multicomponent intervention to prevent physical restraints in nursing homes (IMPRINT): a pragmatic cluster randomized controlled trial. *Int J Nurs Stud*. 2019;96:27–34.
22 Wullschleger A, Vandamme A, Mielau J, et al. Effect of standardized post-coercion review on subjective coercion: results of a randomized-controlled trial. *Eur Psychiatry*. 2021;64(1):e78.
23 Tinland A, Loubière S, Mougeot F, et al. Effect of psychiatric advance directives facilitated by peer workers on compulsory admission among people with mental illness: a randomized clinical trial. *JAMA Psychiatry*. 2022;79:752–759.
24 Abderhalden C, Needham I, Dassen T, Halfens R, Haug HJ, Fischer JE. Structured risk assessment and violence in acute psychiatric wards: randomised controlled trial. *Br J Psychiatry*. 2008;193:44–50.
25 Putkonen A, Kuivalainen S, Louheranta O, et al. Cluster-randomized controlled trial of reducing seclusion and restraint in secured care of men with schizophrenia. *Psychiatr Serv*. 2013;64:850–855.
26 Celofiga A, Kores Plesnicar B, Koprivsek J, et al. Effectiveness of de-escalation in reducing aggression and coercion in acute psychiatric units. A cluster randomized study. *Front Psychiatry*. 2022;13:856153.
27 Bowers L, James K, Quirk A, et al. Reducing conflict and containment rates on acute psychiatric wards: the Safewards cluster randomised controlled trial. *Int J Nurs Stud*. 2015;52:1412–1422.
28 Mullen A, Browne G, Hamilton B, Skinner S, Happell B. Safewards: an integrative review of the literature within inpatient and forensic mental health units. *Int J Ment Health Nurs*. 2022;31:1090–1108.
29 Flammer E, Eisele F, Hirsch S, Steinert T. Increase in coercive measures in psychiatric hospitals in Germany during the COVID-19 pandemic. *PLoS One*. 2022;17:e0264046.
30 Zielasek J, Vrinssen J, Gouzoulis-Mayfrank E. Utilization of inpatient mental health care in the Rhineland during the COVID-19 pandemic. *Front Public Health*. 2021;9:593307.