A Systematic Review of the Effectiveness of Safewards: Has Enthusiasm Exceeded Evidence?

Katie Finch[a], Daniel Lawrence[b,c]*, Marc Williams[a], Andrew Thompson[a] and Christopher Hartwright[a,c]

[a] School of Psychology, Cardiff University, Cardiff, UK

[b] Department of Applied Psychology, Cardiff Metropolitan University, Cardiff, UK

[c] Priory Group, Priory Healthcare, Monmouthshire, UK

*Corresponding author: Daniel Lawrence, Psychology Department, Priory Llanarth Court, Abergavenny, Monmouthshire, NP15 2YD; daniellawrence@priorygroup.com

A Systematic Review of the Effectiveness of Safewards: Has Enthusiasm Exceeded Evidence?

Safewards is intended to be an evidence-based approach to reduce levels of conflict and containment in mental health inpatient settings. A systematic review was carried out to examine whether Safewards is effective in reducing conflict and containment events; and improving ward climate. Searches for articles evaluating the implementation of Safewards was conducted using Psych Info, PubMed, Web of Science, Cochrane Library and CINAHL. Thirteen studies were included for review after applying inclusion and exclusion criteria. The Quality Assessment Tool for Studies with Diverse Designs (QATSDD) was used to assess study quality and the majority of studies (N = 7) were rated as "moderate" quality. Whilst there is evidence to suggest that Safewards is effective for reducing conflict and containment in general mental health services, there is insufficient high-quality empirical evidence to support its effectiveness in settings beyond this. Further research using robust methodological designs with larger, more representative samples is required in order for the effectiveness of Safewards to be established across the range of contexts in which it is currently being applied.

Introduction

The term "conflict" describes behaviours that pose a risk of harm, such as aggression, absconding and self-harm (Bowers, 2014). Prevalence of conflict events, e.g. violence or aggression, is high within health and social care settings, with 200 reported physical assaults estimated on National Health Service (NHS) staff daily in England (Royal College of Nursing, 2018). Mental health services (alongside learning disability services) often report more heightened violence than other care settings, with inpatient mental health services being particularly high risk environments for such behaviours (Cornaggia et al., 2011; National Institute for Health and Care Excellence, 2015; Royal College of Nursing, 2018). Mental health nurses are ten times more likely to be assaulted than general nurses (Parish, 2014).

Conflict places staff and service users at risk of serious harm – including physical harm (Langsrud et al., 2007) and psychological harm, along with trauma and re-trauma (Appleby et al., 2006; Needham et al., 2005; Renwick et al., 2016; Wildgoose et al., 2003). In addition, conflict/violence and aggression disrupt the therapeutic environment and can have negative consequences for therapeutic relationships (Budd, 1999).

The prevalence of conflict behaviours in mental health settings drives the use of "containment". Often termed restrictive or coercive practices in the literature, containment refers to methods through which staff manage conflict behaviours, including restraint, increased observation, seclusion, and security policies (Bowers et al., 2011). This is problematic as a wealth of research evidence demonstrates negative physical and psychological effects of containment methods, including staff and service-users feeling distressed, recalling previous trauma and feeling fearful and angry; it can damage service-user relationships with staff and services and they have been described as incompatible with the basic values of healthcare (Kontio et al., 2014; Steinert et al., 2007; Steinert et al., 2010; Strout, 2010; Stubbs, 2009). Due to these concerns, best-practice guidelines and policy state

that this restrictive approach to managing service users' risks should only be considered as a last resort (Department of Health, 2014), and there has been an international move to reduce the use of them. It has been suggested that evidence-based ways of reducing the need for restrictive practices should be implemented. Several models and programmes have been developed to meet this need.

Safewards (Bowers, 2014) is one such model that has been developed to reduce conflict and containment in mental health inpatient services (see Bowers, 2014 for a comprehensive overview of the model). Safewards consists of ten core interventions that were identified through a rigorous testing and refining process (Bowers et al., 2015). These interventions (outlined in Table 1) were deemed to be the most impactful and easy to implement, resulting in a package of interventions suitable to be evaluated within a randomised control trial. The initial RCT of the model showed promise, with overall rates of containment (coerced medication, seclusion, restraint, special observation etc.) and conflict (i.e. verbal aggression, suicide attempts, alcohol use or attempted absconding) reducing over 31 psychiatric inpatient wards following implementation (Bowers et al., 2015). Due to this success, the model has been recognised within UK guidelines for reducing restrictive practice (Department of Health, 2014; National Institute for Health and Care Excellence (NICE), 2015).

Since then the use of Safewards has grown rapidly with its popularity increasing nationally and internationally as government drivers push for more compassionate and less restrictive ways of working within mental health care. Despite this, no review has comprehensively examined research into the effectiveness of implementing Safewards. Gerdtz et al. (2020) conducted a scoping review of the Safewards literature and aimed to describe the following: Safewards interventions; how Safewards interventions have been implemented in healthcare settings; outcome measures used to evaluate the effectiveness of Safewards; and barriers and enablers to the uptake and sustainability of Safewards. They explicitly did not consider the

effectiveness of the Safewards approach. As such, the need for a systematic review of the literature to examine the effectiveness and impact of the model is warranted. This will allow for the consideration of whether the growing enthusiasm for Safewards exceeds the evidence that supports its use.

[Table 1 near here]

Aims and Objectives

The aim of this systematic review was to collate and examine the literature available to explore the effectiveness of Safewards. A meta-analysis was not possible due to the heterogeneity of the available literature in terms of design, methodology and outcomes used. Therefore, the review could not establish effectiveness by providing a single statistical estimation of the overall effect of the model. Instead the current study presents a narrative synthesis of the data collated systematically, with the aim of minimising bias and improving the reliability of the current evidence base. Whilst the main outcomes of the model are intended to be a reduction in conflict and containment (Bowers et al., 2015), an additional outcome measured within the literature includes the impact on ward climates. Given the negative consequences of conflict and containment on therapeutic environments and interpersonal relationships (Kelly et al., 2015; Renwick et al., 2016; Strout, 2010; Wildgoose et al., 2003) it felt pertinent to include this outcome in order to evaluate the model holistically.

To fulfil the aims of the review, effectiveness was measured with a mixed-methods approach in respect of the following outcomes:

   i.   Does Safewards lead to a reduction in conflict?

ii.   Does Safewards lead to a reduction in containment?

iii.   Does Safewards lead to an improvement in ward climate?

Method

Search Strategy

PsychINFO, Web of Science, Cochrane Library, PubMed and CINAHL were searched for articles. The grey literature, including dissertation, conference and white papers, was also searched using university search engines and dissertation repositories. The term "safeward*" was used as a keyword and topic for all databases and the only keyword used, due to the specific nature of the review and being a protected term in itself in relation to the model. All papers published up until the initial review took place (July 2020) were eligible. Reference lists within the studies were screened and a generic search using Google Scholar was undertaken to ensure completeness. Finally, key authors were contacted to request any further studies that might fit the inclusion criteria. A second search using the same strategy was completed in May 2021 in order to account for any further Safewards publications since the original search. The review was registered with PROSPERO in July 2020.

Study Selection

Titles and abstracts of all records identified were screened for eligibility by the first author then reviewed by the fifth author. Full-text articles were obtained for results that were potentially of interest and screened according to inclusion and exclusion criteria (Table 2). The process of selecting studies is outlined in the Preferred Reporting Items for Systematic Reviews and Meta Analyses flow diagram (PRISMA; Moher et al., 2009). Synthesis without meta-analysis (SWiM) guidelines as described by Campbell et al. (2020) were also followed.

The details pertinent to these guidelines are included throughout the method section. Also relevant to the SWiM guidelines are the results and limitations sections which are included below.

[Table 2 near here]

Search Results

Results of the search are shown in Figure 1. After reviewing a total of 27 papers, 14 were excluded due to not meeting the inclusion criteria. As a result, 13 articles were deemed eligible and included in the review. Reference lists of the included studies were screened and citation searches conducted to ensure relevant articles had not been missed.

[Figure 1 near here]

Assessment of Quality

A quality assessment tool was used to evaluate the methodological rigour of each study. The studies included used a range of research designs, and were both quantitative and qualitative. Therefore, the studies were assessed using a tool designed to evaluate a variety of designs and methods. This also ensured that each study was assessed to the same standard, allowing direct comparisons of quality.

The Quality Assessment Tool for Studies with Diverse Designs (QATSDD: Sirriyeh et al., 2011) was used. The QATSDD consists of 16 items rated using a four-point scale (0 = *Not at all*; 1 =*Very slightly*; 2= *Moderately* and 3 = *Completely*). Of the 16 items, 14 are applicable to quantitative studies and 14 to qualitative. All 16 are applicable to mixed-method designs. Scores are used to calculate a percentage, with higher percentages denoting higher quality

7

papers. The QATSDD has good inter-rater reliability (Ќ =71.5%) and validity for studies with diverse designs (Sirriyeh et al., 2011).

   Assessment of all studies was undertaken by the first author and inter-rater reliability was determined through rating a random selection of studies by the second and third authors. Disagreements were resolved through discussion and abstract screening. Quality of papers was used to guide the review in terms of the implications of each study; and to inform conclusions drawn from outcomes. For this reason, study quality was categorised based on the GRADE approach (GRADE working group 2004). Studies scoring between 0-25% were considered *very low* quality; 26-50% *low* quality; 51-75% *moderate*; and scores of 76-100% were considered to be of *high* quality. This facilitated the direct comparison of study quality and provided a consistent interpretation of QATSDD scores. It should be noted that lower quality studies are more likely to contain systematic errors or higher levels of bias.

Data Extraction

Information from each study was extracted, including: Author(s); country; setting; design; total sample size; sample characteristics; outcomes and outcome measure tool(s); primary findings; quality assessment scores and quality categorisation. The information was then reported in three standardised tables by the first author and reviewed by the second author. The studies were grouped in tables based on the outcomes they reported relative to the aims of this review. Specifically, studies reporting on the effect of Safewards on conflict were included in Table 3. Those reporting on containment in Table 4; and those on ward climate in Table 5. Naturally, some studies focused on more than one of these outcomes and where this was the case they were included in multiple tables. Studies were arranged in tables based on their quality score, where higher priority was given to studies of higher quality as defined by QATSDD criteria. Initially, all studies were recorded in one large table but the authors felt

that it was better to divide them into tables based on outcomes in order to display the results in a manner that would more clearly address the objectives of the review. Due to the diverse nature of the studies including the diversity in design, outcomes, and the way in which results were reported it was not possible to standardise and synthesise results further than what is reported. However, the current authors did use p values as the standardised metric for results pertaining to the effectiveness in reducing conflict and containment. Effect sizes were reported if available. The majority of studies concerned with ward climate used the same measure to evaluate this and typically reported on differences in mean scores. A number of studies reported qualitative data regarding perceived effectiveness of the model.

[Table 3 near here]

[Table 4 near here]

[Table 5 near here]

Results

Quality of Included Studies

Quality assessment scores can be seen in Table 3. QATSDD scores were rated out of a total of 48 for studies of mixed design. For either quantitative or qualitative studies, the maximum score was 42. The included mixed methods studies obtained scores between 14- 30 with a mean score of 23 (5.90). For those with either quantitative or qualitative methods, scores ranged from 26 to 34 with a mean score of 30.17(4.02).

Only one study (Bowers et al., 2015) adopted a Randomised Control Trial (RCT) design. Most studies failed to recruit a representative sample of a reasonable size and only 25% obtained a moderate rating for this item. Whilst all studies chose methods of data collection and analysis appropriate to their research question, many failed to report a justification for

9

their design. Only 23% scored two or above for QATSDD criterion "Statistical assessment of reliability and validity of measurement tool(s) (Quantitative only)". Positively, most studies provided a rationale for their method of analysis with 77% of studies scoring two or above on this item. In addition, 69% of studies scored two or more for "Evidence of sample size considered in terms of analysis" which is important for ensuring statistical rigour for studies of a particular type of design (e.g. RCT).

Numerous studies failed to provide a clear enough description of the setting (Cabral & Carthy, 2017; Stensgaard et al., 2018), their data collection procedure (Baumgardt et al., 2019; Cabral & Carthy, 2017; Lickiewicz et al., 2020; Price et al., 2016) and recruitment process (Baumgardt et al., 2019; Cabral & Carthy, 2017; Davies et al., 2020; Fletcher et al., 2019a; 2019b; Price et al., 2016; Stensgaard et al., 2018) to allow for the research to be replicated.

Strengths across studies included "Statement of aims/objectives in main body of report", "Fit between stated research question and method of data collection" and "Fit between research question and method of analysis", with 100% of studies scoring two (*moderately*) or three (*completely*) on these criteria. Over 90% scored two or above for having an "Explicit theoretical framework".

An additional strength is the use of shared target outcomes across studies which allowed for direct comparisons between them. However, there were discrepancies between some of the studies with regards to their monitoring of containment methods. Lickiewicz et al. (2020) chose to compare the frequency of mechanical restraint, whereas Stensgaard et al. (2018) compared frequency of coercive measures (which included mechanical restraint as well as forced sedation). Some collected information on "seclusion" only (Fletcher et al., 2017).

Safewards Interventions and Mode of Implementation

Some studies used more comprehensive implementation strategies than others. This is important when considering the results of studies and has implications for staff exposure and fidelity to the model. Of the 13 studies included for review, 12 implemented all ten Safewards interventions. One study (Lickiewicz et al, 2020) evaluated three interventions only: "Positive words"; "reassurance"; and "clear mutual expectations". Most (Baumgardt et al., 2019; Davies et al., 2020; Dickens et al., 2020; Fletcher et al., 2017; Hottinen et al., 2019; Lickiewicz et al., 2020; Maguire et al., 2018; Stensgaard et al., 2018) provided some form of staff education/training including workshops, 'train the trainer' sessions, staff meetings or education meetings. Three studies used steering groups (Baumgardt et al., 2019; Dickens et al., 2020; Maguire et al., 2018), 1 reported a local working party (Maguire et al., 2018) and four identified 'Safewards champions' on wards or recruited specific staff to take responsibility for implementing interventions (Baumgardt et al., 2019; Davies et al., 2020; Dickens et al., 2020; Hottinen et al., 2019).

Two studies (Baumgardt et al., 2019; Dickens et al., 2020) used a number of the above methods for their implementation strategy, including a steering group, training workshops, Safewards champions and responsible individuals/groups. Baumgardt et al. (2019) in particular provided the most detail and illustrated the most well-designed strategy for implementation of the model. This meant that it was also the most replicable. By contrast, the only RCT of Safewards (Bowers et al., 2015), failed to provide a detailed description of the process by which staff/key personnel were trained in the interventions.

Impact on Conflict

Seven of the 13 studies reported quantitative outcomes relating to conflict. Four used the Patient-Staff Conflict Checklist (PSCC) or Patient-Staff Conflict Checklist- Shift Report

(PSCSR; Bowers et al., 2015; Davies et al., 2020; Dickens et al., 2020; Price et al., 2016), which examines 22 conflict items within broader categories of "aggression", "self-harm", "general rule breaking", "substance use", "absconding" and "medication related". One study, Maguire et al. (2018) measured "aggressive incidents" (physical, verbal and towards property) alongside "attempted absconding", "affected by drugs/alcohol", "self-harm" and "medication refusal". Fletcher et al. (2019a) did not look at conflict events per se and instead used an online survey with both quantitative and qualitative questions regarding the impact of Safewards from the perspective of service users. Fletcher et al. (2019b) used similar approach but sought to determine the impact of Safewards from the perspective of staff.

All five studies concerned with conflict events reported a reduction in them following the implementation of Safewards, 3 of which reported statistically significant reductions (Bowers et al., 2015; Davies et al., 2020; Dickens et al., 2020). The highest quality study (Bowers et al., 2015) reported a significant reduction in the number of conflict events. This study found that Safewards was more effective in reducing conflict than the control intervention, although some decrease in conflict was seen in control wards. Effect sizes were not reported, meaning the magnitude of this difference could not be concluded. Dickens et al. (2020; also of high quality) reported a significant reduction in conflict events between the baseline and implementation phases of their study; and between the baseline and outcome phases; but not between the implementation and outcome phases. The same findings were observed when the authors looked at the most serious conflict events (physical aggression). Davies et al. (2020) reported significant reductions for some, but not all, conflict events including: Aggression; absconding; and medication related behaviours. Nine of the 31 sub-items within the PSCSR showed significant reductions. Whilst of moderate quality, difficulties in the study relating to sample consistency were noted. This impacts on the individual characteristics and

presentations exposed to the interventions as well as the overall exposure rate, which should be considered when interpreting results.

Price et al. (2016) and Maguire et al. (2018) both found a reduction of conflict but this did not reach statistical significance in the former and statistical analyses of significance were not used in the latter study. The use of a service-evaluation design as well as low adherence rates in Price et al. (2016) poses significant limitations on how these results can be interpreted. Ultimately, the degree to which interventions contributed to the reported reductions in conflict was unclear. This is also true of Maguire et al. (2018) in that it cannot be concluded whether reductions were meaningful or significant, nor if they were the result of the interventions themselves.

Fletcher et al. (2019a) found that service users felt Safewards had a positive impact on physical and verbal conflict, where 25% of participants reported that Safewards 'usually' or 'always' helped to resolve physical and verbal conflict. No analysis of significance was conducted. Fletcher et al. (2019b) reported similar findings where staff were of the view that Safewards impacted on physical and verbal conflict, with 45% and 55%, respectively, reporting that Safewards usually or always had a favourable impact. Staff were less confident of the positive impact on other forms of conflict such as absconding and property damage, reporting that Safewards usually or always positively impacted 30% and 35%, respectively.

In terms of qualitative data on conflict, participants in the study by Maguire et al. (2018) were of the view that Safewards improved conflict rates and that there were less incidents following the introduction of the intervention. Conversely, Price et al. (2016) summarised that the dominant view among the staff participants in their study was that, although the interventions included many good ideas, they did not deem them to be effective for conflict rates on the male wards in their forensic mental health service.

Impact on Containment

Nine studies reported quantitative outcomes relating to containment use. As noted, the definition of "containment" and subsequent outcomes varied across studies. Key events measured include those within the PSCC and PSCSR: "Given PRN medication (psychotropic)"; "Given IM medication (enforced)"; "Sent to PICU or ICA"; "Seclusion"; "Special observation (intermittent)"; "Special observation (continuous)": "Show of force"; "Manually restrained" and "Time out". Additional events measured included: "Mechanical restraint"; "limitation to free movement"; "forced medication" and "forced sedation". Rates reported included both individual events measured as well as combinations or overall totals of containment. Variations in how containment was measured occurred in part due to the range of countries where studies were conducted. Ways of managing conflict differ internationally due to legal, organisational and cultural reasons.

The use of the PSCC or PSCSR to measure containment was seen in four studies (Bowers et al., 2015; Davies et al., 2020; Dickens et al., 2020; Price et al., 2016). Despite using the same checklists, variations were seen across the studies with two using eight descriptions of containment (Bowers et al., 2015; Price et al., 2016) and Davies et al (2020) reporting only six. This was due to 'blending' of categories such as "PRN medication" and "intramuscular medication (enforced)" and "special observations (continuous)" and "special observations (intermittent)". Dickens et al. (2020) measured and distinguished between overall containment rates and rates of "highly coercive" containment (i.e. seclusion, restraint and forced IM). Where studies did not use the PSCC, measurements of containment included accessing routine incident data (Fletcher et al., 2017; Lickiewicz et al., 2020; Maguire et al., 2018; Stensgaard et al., 2018) or collecting specific data for containment incidents (frequency and duration; Baumgardt et al., 2019).

Eight studies reported reductions in containment events, six of which reported statistically significant findings (Baumgardt et al., 2019; Bowers et al., 2015; Davies et al., 2020; Dickens et al., 2020; Lickiewicz et al., 2020; Stensgaard et al., 2018). All studies were of moderate or high quality. Fletcher et al. (2017) and Price et al. (2016) both reported non-significant reductions in containment and were of high and low quality respectively. One study (Maguire et al., 2018), reported no change in the rates of seclusion and increases in physical and mechanical restraint. The study was of moderate quality; however, it is acknowledged that sample size was small and existing levels of conflict and containment were already low, which meant meaningful statistical analysis was not feasible. Whilst frequency of containment was the primary outcome, one study did also demonstrate a significant reduction in duration of containment measures overall (Baumgardt et al., 2019), however the reported effect size indicated a small effect of the interventions.

Maguire et al. (2018) and Price et al. (2016) both reported qualitative data regarding the impact of Safewards from the perception of staff. Participants in the study by Maguire et al. (2018) were of the view that Safewards improved containment rates. On the other hand, Price et al. (2016) reported that the consensus view among the staff in their study was that Safewards interventions were not effective for containment rates on the male wards in their forensic mental health service.

Impact on Ward Atmosphere or Climate

Three studies used the Essen Climate Evaluation Schema Questionnaire (EssenCES; Schalast et al., 2008) to evaluate change in ward climate following Safewards implementation. The EssenCES is a 15-item questionnaire designed to measure social climate using three sub-scales that evaluate 'experienced safety', 'patient cohesion' and 'therapeutic hold' (rating of relational security).

Hottinen et al. (2019) found that staff 'experienced safety' and service user 'therapeutic hold' scores improved significantly from baseline to follow-up. Service user 'patient cohesion' scores also improved significantly when compared with baseline measures. Maguire et al. (2018) found a significant increase in staff perceptions of 'patient cohesion' and staff 'experienced safety'. Cabral and Carthy (2017) reported an improvement of mean scores across all three sub-scales following implementation of Safewards. However, this study was low quality and had a limited sample size. Further statistical analyses were not conducted and the significance of these changes cannot be inferred.

As opposed to the EssenCES, one study (Dickens et al., 2020) used the Violence Prevention Climate-14 (VPC-14; Hallett et al., 2018) to measure the ward climate. They reported that violence prevention climate ratings did not change following the implementation of Safewards. It is noteworthy that this was the highest quality study that reported on ward climate.

Discussion

The aim of this review was to identify outcome studies reporting implementation of Safewards, evaluate their scientific rigour and synthesise the data in order to investigate the effectiveness of the Safewards model. Thirteen studies were identified with considerable variation in methodological designs. Quality of the studies also varied, with only four of the studies being assessed as "high" quality, seven "moderate" quality and two "low" quality. Only one study utilised a RCT design allowing a more robust evaluation of intervention outcomes, although this still presented with some methodological flaws. That being said this particular study (Bowers et al., 2015) along with a number of other high quality studies provide support for the effectiveness of Safewards in reducing conflict and containment.

However, based on the findings of the current review, the authors are of the opinion that this effect is currently limited to general mental health settings as this is the type of setting the high majority of papers (n=9) in the review were concerned with. The aims of this review will be used to structure the discussion of current evidence for the model and to consider ongoing clinical and research implications.

Does Safewards Lead to a Reduction in Conflict?

Safewards interventions were found to reduce levels of conflict (Bowers et al., 2015; Davies et al., 2020; Dickens et al., 2020; Maguire et al., 2018; Price et al., 2016), however only three studies reported statistically significant changes. Only one study used a strong methodological design (RCT; Bowers et al., 2015) but a number of methodological limitations need consideration when interpreting its findings. Mustafa (2015) noted that the "true" degree of blinding of participants in the study by Bowers et al. (2015) cannot be ascertained given the likely differences between the control and experimental conditions as well as the movement of staff between wards. It is also unclear how significant findings were reported despite low rates of fidelity and exposure to the Safewards model. It was not possible to determine whether other factors confounded the findings, or whether Safewards was effective despite low adherence. Given the nature of the control condition (a staff physical health intervention), the low adherence to the Safewards model also suggests that it is possible that the reductions in conflict and containment in the intervention wards were not as a result of Safewards specifically, but occurred as a result of the implementation of an intervention that is designed to address conflict and containment. It might have been that any model designed for this purpose would have resulted in better results than the control condition. Further research that considers these issues and the importance of fidelity to the Safewards model is required.

Across all of the included studies, the exact nature of "conflict events" varied and as such there was little consistency between outcome measures. Davies et al. (2020) found that Safewards had an effect on some conflict events but not all, making it difficult to interpret the intervention's overall effectiveness. It was unclear whether findings were due to better implementation of some interventions over others, or if interventions differ in their targeting of different types of conflict. This lack of clarity and consistency impedes the ability to establish effectiveness with regards to specific conflict-related outcomes.

Processes of implementation varied too, with no two studies following the same strategy or number of resources/procedures to establish interventions. This makes it difficult to draw conclusions as to whether external factors (e.g. staff buy-in, management/organisational support) had any impact on the ability of Safewards to reduce conflict. These findings replicate previous reviews concerned with conflict/containment where the equivocal nature of the research (Goulet et al., 2017; Woods & Ashley, 2007) and the presence of complex factors at play drive recommendations for multifaceted/multi-layered interventions (Duxbury, 2015). Despite this, this review does indicate that the Safewards interventions are effective for reducing conflict events. However, as is noted above, this effect is limited to the use of Safewards in general mental health settings and it is currently not possible to state that Safewards is effective in reducing conflict in settings beyond this. For example, only three studies in the current review looked at the implementation of Safewards in secure/forensic mental health settings, and only two of which were concerned with conflict. Maguire et al. (2018) was assessed as being of moderate quality and Price et al. (2016) of low quality. Both reported reductions in conflict but the former did not test for statistical significance and the reductions failed to reach statistical significance in the latter. Only one study was concerned with the use of Safewards with an Intellectual Disability population and reported significant reductions in conflict (Davies et al., 2019). This study was of moderate quality. As such, it is

concluded that further, high quality evidence is required before Safewards can be deemed as an effective intervention for reducing conflict with people who use inpatient services beyond general mental health settings. This is important as individuals in different settings have been found to differ in terms of their presenting needs and risks (e.g. Timmerman & Emmelkamp, 2001).

Does Safewards Lead to a Reduction in Containment?

Impact on containment was evaluated more frequently than conflict. This is surprising given the close relationship between them (Bowers & Allan, 2006) and that the main aim of Safewards is to reduce both. Nine studies evaluated containment (Baumgardt et al., 2019; Bowers et al., 2015; Davies et al., 2020; Dickens et al., 2020; Fletcher et al., 2017; Lickiewicz et al., 2020; Maguire et al., 2018; Price et al., 2016; Stensgaard et al., 2018). Eight reported reductions in containment, six of which reported significant reductions. All studies were of moderate or high quality. Similar methodological limitations to those listed in the preceding section were present. Again, only one study adopted a RCT design and multiple studies had small sample sizes, limiting the generalisability of findings. The findings of the current review correspond with other reviews of restraint reduction programmes (Goulet et al., 2017; Scanlan., 2010) that demonstrate overall positive changes following intervention implementation but also note the limited quality and scarcity of empirical research.

The inability to determine the most successful of the ten interventions; or whether individual interventions are more effective than combined are key difficulties when interpreting outcomes of Safewards. Bowers et al. (2014) recognised this when creating the model. Whilst we therefore cannot conclude why there were improvements, the findings of this review support the use of Safewards as a method to reduce rates of containment. However, as noted in the above section, and for the same reasons, we are of the view that this

effect is currently limited to general mental health inpatient settings and further research is required before Safewards can be deemed to be an effective way of reducing containment in settings beyond this. Again, further research that takes this in to consideration is important as containment rates have found to differ across services where they are higher in forensic compared to general mental health services for example (e.g. Flammer et al., 2020).

Does Safewards lead to an Improvement in Ward Climate?

Improvement in ward climate was reported in three studies (Cabral & Carthy, 2017; Hottinen et al., 2019; Maguire et al., 2018), two of which reported significant changes. The consistent use of a validated and reliable measure across these studies improves the scientific robustness of each study. Methodological flaws described above were present in these studies, including small sample sizes, limiting the ecological validity and generalisability of findings. One study, Cabral and Carthy (2017), was also assessed as low quality with low internal validity. The highest quality study that considered ward climate (Dickens et al., 2020) found no change in ratings after Safewards was implemented.

The findings on ward climate are at this stage modest but we can tentatively suggest that Safewards can have a positive impact on ward climate, which may be an obvious consequence of reductions in conflict and containment rates. Interestingly, Maguire et al. (2018) did not see any reductions in containment despite improved ward climate. More research that focuses on ward climate is required including research that will help to understand the mechanisms through which Safewards can improve ward climate.

Clinical Implications and Recommendations

The findings of this review suggest that Safewards is an effective way to reduce conflict and containment in general mental health inpatient services, and it may also be a helpful approach

to improve the overall climate of a ward. This provides useful information for clinicians and healthcare providers alike, particularly within the mental-health field and where conflict and containment events are prevalent. The review suggests that Safewards may provide a helpful alternative to restrictive practices for reducing conflict events in such settings. This is important to consider, given the negative consequences of restrictive practices on both patients and staff (Appleby et al., 2006; Needham et al., 2005) and the risk of traumatisation within inpatient settings in particular (Cusack et al., 2018).

By introducing Safewards and adapting existing ways of working as well as considering the role staff play in preventing conflict and containment, reductions in incidents can be achieved. Clinically this has implications for how patient care is delivered, how staff are trained to work in inpatient settings, and how staff respond to challenges on wards. As well as service user wellbeing, this has clear implications for staff wellbeing and professional quality of life. Considering the problems associated with restrictive practices and current government advice and policy, ongoing efforts to reduce their use are required. This review suggests that Safewards is an effective way to achieve this aim in general mental health services.

As well as the main findings, some of the included studies (e.g. Price et al., 2016) reported significant barriers to the implementation of Safewards, that were difficult to overcome and ultimately contributed to the lack of success observed. If services are to be successful in implementing Safewards in the future, sufficient thought and planning will be required to address these issues. Limited or poorly planned implementation strategies, for example as a result of limited resources dedicated to the model, are likely to inhibit the success of Safewards.

It is however important to note that at this stage, the above listed implications of this review and the potential benefits of Safewards are mostly limited to general mental health

settings and further research is required to determine the clinical benefits of using Safewards in settings beyond this.

Research Implications

A better understanding of the effectiveness of the Safewards model and its impact on both staff and service-users requires addressing through future research. Currently there is limited available evidence to support the use of Safewards as a strategy to reduce conflict and containment in settings beyond general mental health inpatient services. This may reflect the fact that model is currently in its infancy. As commented upon earlier in the current paper, whilst there is a lack of Safewards related research publications available, there is no such lack of enthusiasm for the model and it has been applied in a wide range of settings internationally, including those other than general mental health wards. Practitioners and academics are encouraged to continue to evaluate and publish such work so that the evidence base supporting Safewards can continue to grow.

As well as the quantity of Safewards research, it is crucial that future studies are of sufficient quality to contribute meaningfully to the evidence-base by using larger samples, more robust methodological designs over longer periods and with clear processes to reduce risk of bias and improve replicability. This is particularly necessary for research conducted in settings beyond general mental health settings (e.g. forensic/secure; and intellectual disability services). One primary recommendation for future research is to consider how to embed Safewards holistically and at a wider systemic level, to increase buy-in and motivation from organisations and staff in order to improve the chances of success for the model in the settings where it is being applied. It is noteworthy that the current review found a number of qualitative studies that focused on staff and service users' acceptance of the model, their views on the usefulness of it and perceptions on barriers to implementation. These studies

were excluded from our review because they were beyond the scope of what we aimed to address. However, future research may wish to consider this area, as a metasynthesis of these qualitative studies would likely provide information useful to the processes required to successfully implement and maintain the Safewards interventions. Future research should also attempt to employ outcome measures that improve the consistency of reporting across studies and that are in themselves valid and reliable. This will in turn improve the quality of research and strength of findings.

Finally, future research should consider establishing which of the interventions are most impactful, whether all interventions are needed to promote positive change and if not, which combination of interventions is most effective. As seen in this review, utilisation of just three of the 10 interventions can produce significant results (Lickiewicz et al., 2020). Given the ongoing pressures and demands of staff within healthcare settings, making new models and interventions easier to implement and minimising the amount of new knowledge and skills required to do so, may have a beneficial effect on staff uptake and adherence resulting in more controlled and reliable research.

Strengths and Limitations

The use of a quality assessment tool (QATSDD, Sirriyeh et al., 2011) to examine the rigour of included studies and to inform conclusions drawn from the research is a strength of the current paper. The authors note however that this tool has been criticised in the past for lack of explicit language and examples to guide assessment (Fenton et al., 2015). A further strength is that, due to the protected nature of the name, Safewards was neatly defined and thus we can be confident that all studies evaluating Safewards that met the inclusion criteria have been reviewed. This also ensured the consistency of interventions applied across studies.

Several limitations to this review should be noted. First, a formal reliability check of quality assessing was only done partially, whereby a random sample of the papers were assessed independently by the second and third authors. It would have been more robust for all studies to have been quality assessed independently by two of the researchers.

In addition, there were few high-quality studies available to be included, meaning this paper can only synthesise data from studies that require improvement to methodological design and scientific rigour. Only one study reported effect size which was small (Baumgardt et al., 2019), meaning that the actual magnitude of effect the interventions had on outcomes (and therefore degree of effectiveness) cannot be determined. It is also acknowledged that, due to the variation in outcomes reported, diverse study designs and lack of RCTs, it was not possible to complete a meta-analysis or standardise/synthesise the results any further than what has been reported. Conclusions drawn about the effectiveness of Safewards is limited without a precise, statistical estimate of the overall effect of the model on the reported outcomes. A more pragmatic approach and possible random-effects model for a meta-analysis would significantly improve the statistical strength of conclusions made regarding the effectiveness of Safewards. However, despite this limitation, the authors are of the view that the collation of all Safewards studies concerned with the effectiveness of the model into one coherent review provides a good foundation from which to develop the Safewards evidence base. Finally, it is important to recognise the limited generalisability of the findings of this review, given the small number of studies and limited sample sizes within those included.

Conclusion

This paper was the first to systematically review the evidence for the Safewards model and examine the empirical basis for its effectiveness in reducing conflict and containment in

healthcare settings. Enthusiasm for Safewards has been observed where it has been employed in a wide range of settings internationally. Whilst this review indicates that there is good evidence to support its use in general mental health services, it is not yet possible to say that it is an effective model for reducing conflict and containment beyond this context. Further robust research is imperative in order to determine effectiveness in other clinical settings and to continue contributions to the model's growing evidence base.

Disclosure Statement

There are no interests to declare.

References

Appleby, L., Shaw, J., Japur, N., Windfuhr, K., Ahston, A., Swinson, N., & While, D. (2006). *Avoidable Deaths: Five Year Report of the National Confidential Inquiry into Suicide and Homicide by People with Mental Illness*. University of Manchester, Manchester.

Baumgardt, J., Jäckel, D., Helber-Böhlen, H., Stiehm, N., Morgenstern, K., Voigt, A., ... & Schulz, M. (2019). Preventing and reducing coercive measures—an evaluation of the implementation of the Safewards model in two locked wards in Germany. *Frontiers in Psychiatry*, *10*, 340. DOI: 10.3389/fpsyt.2019.00340.

Bowers, L., & Allan, T. (2006). The Attitude to Personality Disorder Questionnaire: psychometric properties and results. *Journal of Personality Disorders*, *20*(3), 281-293. DOI: 10.1521/pedi.2006.20.3.281.

Bowers, L., Stewart, D., Papadopoulos, C., Dack, C., Ross, J., Khanom, H., & Jeffery, D. (2011). Inpatient violence and aggression: a literature report from the Conflict and Containment Reduction Research Programme, Kings College, London, available at: www. kcl. ac. uk/iop/depts/hspr/research/ciemh/mhn/projects/litreview/LitRevAgg. pdf.

Bowers, L. (2014). Safewards: a new model of conflict and containment on psychiatric wards. *Journal of Psychiatric and Mental Health Nursing*, *21*(6), 499–508. DOI: 10.1111/jpm.12129.

Bowers, L., Alexander, J., Bilgin, H., Botha, M., Dack, C., James, K., ... & Papadopoulos, C. (2014). Safewards: the empirical basis of the model and a critical appraisal. *Journal of Psychiatric and Mental Health Nursing*, *21*(4), 354-364. DOI: 10.1111/jpm.12085

Bowers, L., James, K., Quirk, A., Simpson, A., Stewart, D., & Hodsoll, J. (2015). Reducing conflict and containment rates on acute psychiatric wards: The Safewards cluster randomised controlled trial. *International Journal of Nursing Studies*, *52*(9), 1412-1422. DOI: 10.1016/j.ijnurstu.2015.05.001

Budd, T. (1999). *Violence at work: Findings from the British crime survey*. London: Home Office.

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., ... & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *British Medical Journal*, *368,* l6890. DOI: 10.1136/bmj.l6890

Cornaggia, C. M., Beghi, M., Pavone, F., & Barale, F. (2011). Aggression in psychiatry

wards: a systematic review. *Psychiatry Research*, *189*(1), 10-20. DOI:

10.1016/j.psychres.2010.12.024.

Cusack, P., Cusack, F. P., McAndrew, S., McKeown, M., & Duxbury, J. (2018). An

integrative review exploring the physical and psychological harm inherent in using

restraint in mental health inpatient settings. *International Journal of Mental Health

Nursing*, *27*(3), 1162–1176. DOI: 10.1111/inm.12432.

Davies, B., Silver, J., Josham, S., Grist, E., Jones, L., Francis, N., ... & Gwatkin, A. (2020).

An evaluation of the implementation of Safewards on an assessment and treatment unit

for people with an intellectual disability. *Journal of Intellectual Disabilities.* DOI:

10.1177/1744629520901637.

Department of Health. (2014). Positive and Proactive Care: reducing the need for restrictive

interventions. London. Retrieved from

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment

_data/file/300293/JRA_DoH_Guidance_on_RP_web_accessible.pdf

Dickens, G. L., Tabvuma, T., Frost, S. A., & SWSLHD Safewards Steering Group. (2020).

Safewards: Changes in conflict, containment, and violence prevention climate during

implementation. *International Journal of Mental Health Nursing*, *29*(6), 1230-1240.

DOI: 10.1111/inm.12762

Duxbury, J. A. (2015). Editorial: Minimizing the use of coercive practices in mental health: the perfect storm. *Journal of Psychiatric and Mental Health Nursing, 22*, 89-91. DOI: 10.1111/jpm.12206.

Fenton, L., Lauckner, H., & Gilbert, R. (2015). The QATSDD critical appraisal tool: comments and critiques. *Journal of Evaluation in Clinical Practice*, *21*(6), 1125-1128. DOI: 10.1111/jep.12487.

Flammer, E., Frank, U., & Steinert, T. (2020). Freedom restrictive coercive measures in forensic psychiatry. *Frontiers in psychiatry*, *11*, 146. DOI: 10.3389/fpsyt.2020.00146

Fletcher, J., Buchanan-Hagen, S., Brophy, L. M., Kinner, S., & Hamilton, B. (2019a). Consumer perspectives of safewards impact in acute inpatient mental health wards in Victoria, Australia. *Frontiers in Psychiatry*, *10*, 461. DOI: 10.3389/fpsyt.2019.00461.

Fletcher, J., Hamilton, B., Kinner, S. A., & Brophy, L. M. (2019b). Safewards impact in inpatient mental health units in Victoria Australia: Staff perspectives. *Frontiers in Psychiatry*, *10*, 462. DOI: 10.3389/fpsyt.2019.00462.

Fletcher, J., Spittal, M., Brophy, L., Tibble, H., Kinner, S., Elsom, S., & Hamilton, B. (2017). Outcomes of the Victorian Safewards trial in 13 wards: Impact on seclusion rates and fidelity measurement. *International Journal of Mental Health Nursing*, *26*(5), 461-471. DOI: 10.1111/inm.12380.

Gerdtz, M., Daniel, C., Jarden, R., & Kapp, S. (2020). Use of the Safewards Model in

healthcare services: a mixed-method scoping review protocol. *British Medical Journal Open*, *10*(12), e039109. DOI: 10.1136/bmjopen-2020-039109.

Gooding, P., McSherry, B., Roper, C., & Grey, F. (2018). *Alternatives to coercion in mental health settings: A literature review*. Melbourne: Melbourne Social Equity Institute, University of Melbourne.

Goulet, M. H., Larue, C., & Dumais, A. (2017). Evaluation of seclusion and restraint reduction programs in mental health: a systematic review. *Aggression and Violent Behavior*, *34*, 139-146. DOI: 10.1016/j.avb.2017.01.019.

GRADE Working Group. (2004). Grading quality of evidence and strength of recommendations. *British Medical Journal, 328,* 1490-1498. DOI: 10.1136/bmj.328.7454.1490.

Hallett, N., Huber, J., Sixsmith, J. & Dickens, G. L. (2018). Measuring the violence prevention climate: Development and evaluation of the VPC-14. *International Journal of Nursing Studies, 88*, 97–103. DOI: 10.1016/j.ijnurstu.2018.09.002

Hottinen, A., Rytilä-Manninen, M., Laurén, J., Autio, S., Laiho, T., & Lindberg, N. (2020). Impact of the implementation of the safewards model on the social climate on adolescent psychiatric wards. *International Journal of Mental Health Nursing*, *29*(3), 399-405. DOI: 10.1111/inm.12674.

Langsrud, K., Linaker, O. M., & Morken, G. (2007). Staff injuries after patient–staff

    incidences in psychiatric acute wards. *Nordic Journal of Psychiatry*, *61*(2), 121-125.

    DOI: 10.1080/08039480701226104.

Lickiewicz, J., Adamczyk, N., Hughes, P. P., Jagielski, P., Stawarz, B., & Makara-

    Studzińska, M. (2020). Reducing aggression in psychiatric wards using Safewards-A

    Polish study. *Perspectives in Psychiatric Care*, *57*(1), 50-55. DOI: 10.1111/12523

Maguire, T., Ryan, J., Fullam, R., & McKenna, B. (2018). Evaluating the introduction of the

    safewards model to a medium-to long-term forensic mental health ward. *Journal of*

    Forensic Nursing, *14*(4), 214-222. DOI: 10.1097/JFN.0000000000000215.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred

    reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS*

    Medicine, *6*(7), e1000097. DOI: 10.1371/journal.pmed.1000097.

National Institute for Health and Care Excellence. (2015). Violence and aggression: short-

    term management in mental health, health and community settings. (NG10). Retrieved

    from https://www.nice.org.uk/guidance/ng10.

Needham, I. A. C. H., Abderhalden, C., Halfens, R. J. G., Dassen, T., Haug, H. J., & Fischer,

    J. E. (2005). The effect of a training course in aggression management on mental health

    nurses' perceptions of aggression: a cluster randomised controlled trial. *International*

    Journal of Nursing Studies, *42*(6), 649-655. DOI: 10.1016/j.ijnurstu.2004.10.003.

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative

research. *Evidence-Based Nursing*, *18*(2), 34-35. DOI: 10.1136/eb-2015-102054.

Parish, C. (2014). A report that hits home. *Mental Health Practice*, *18*(4), 5. DOI:

10.7748/mhp.18.4.5.s1.

Price, O., Burbery, P., Leonard, S. J., & Doyle, M. (2016). Evaluation of safewards in

forensic mental health. *Mental Health Practice*, *19*(8). DOI: 10.7748/mhp.19.8.14.s17.

Renwick, L., Stewart, D., Richardson, M., Lavelle, M., James, K., Hardy, C., ... & Bowers, L.

(2016). Aggression on inpatient units: clinical characteristics and

consequences. *International Journal of Mental Health Nursing*, *25*(4), 308-318. DOI:

10.1111/inm.12191.

Royal College of Nursing (2018). Violence and aggression in the NHS: Estimating the size

and the impact of the problem- an interim report. London. Retrieved from

https://www.rcn.org.uk/professional-development/publications/pub-007301.

Scanlan, J. N. (2010). Interventions to reduce the use of seclusion and restraint in inpatient

psychiatric settings: what we know so far, a review of the literature. *International*

Journal of Social Psychiatry, *56*(4), 412-423. DOI: 10.1177/0020764009106630.

Schalast, N., Redies, M., Collins, M., Stacey, J., & Howells, K. (2008). EssenCES, a short

    questionnaire for assessing the social climate of forensic psychiatric wards. *Criminal*

    Behaviour and Mental Health, *18*(1), 49-58. DOI: 10.1002/cbm.677.

Sirriyeh, R., Lawton, R., Gardner, P., & Armitage, G. (2012). Reviewing studies with diverse

    designs: the development and evaluation of a new tool. *Journal of Evaluation in Clinical*

    Practice, *18*(4), 746-752. DOI: 10.1111/j.1365-2753.2011.01662.x.

Steinert, T., Bergbauer, G., Schmid, P., & Gebhardt, R. P. (2007). Seclusion and restraint in

    patients with schizophrenia: clinical and biographical correlates. *The Journal of Nervous*

    and Mental Disease, *195*(6), 492-496. DOI: 10.1097/NMD.0b013e3180302af6.

Steinert, T., Lepping, P., Bernhardsgrütter, R., Conca, A., Hatling, T., Janssen, W., ... &

    Whittington, R. (2010). Incidence of seclusion and restraint in psychiatric hospitals: a

    literature review and survey of international trends. *Social Psychiatry and Psychiatric*

    Epidemiology, *45*(9), 889-897. DOI: 10.1007/s00127-009-0132-3.

Stensgaard, L., Andersen, M. K., Nordentoft, M., & Hjorthøj, C. (2018). Implementation of

    the safewards model to reduce the use of coercive measures in adult psychiatric inpatient

    units: an interrupted time-series analysis. *Journal of Psychiatric Research*, *105*, 147-

    152. DOI: 10.1016/j.jpsychires.2018.08.026.

Stubbs, B. (2009). The manual handling of the aggressive patient: a review of the risk of

    injury to nurses. *Journal of Psychiatric and Mental Health Nursing*, *16*(4), 395-400.

    DOI: 10.1111/j.1365-2850.2008.01354.x.

Strout, T. D. (2010). Perspectives on the experience of being physically restrained: An

    integrative review of the qualitative literature. *International Journal of Mental Health*

    Nursing, *19*(6), 416-427. DOI: 10.1111/j.1447-0349.2010.00694.x.

Timmerman, I. G., & Emmelkamp, P. M. (2001). The prevalence and comorbidity of Axis I and

    Axis II pathology in a group of forensic patients. *International Journal of Offender*

    Therapy and Comparative Criminology, *45*(2), 198-213. DOI:

    10.1177/0306624X01452006

Wildgoose, J., Briscoe, M., & Lloyd, K. (2003). Psychological and emotional problems in

    staff following assaults by patients. *Psychiatric Bulletin*, *27*(8), 295-297. DOI:

    10.1192/pb.27.8.295.

Woods, P., & Ashley, C. (2007). Violence and aggression: A literature review. *Journal of*

    Psychiatric and Mental Health Nursing, *14*(7), 652-660. DOI: 10.1111/j.1365-

    2850.2007.01149.x.