International Journal of Nursing Studies 121 (2021) 103984

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns

Interventions to mitigate moral distress: A systematic review of the literature

Georgina Morley[a,*], Rosemary Field[b], Cristie Cole Horsburgh[c], Christian Burchill[d,**]

[a] Center for Bioethics and Stanley S. Zielony Institute for Nursing Excellence, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH, USA
[b] Stanley S. Zielony Institute for Nursing Excellence, Cleveland Clinic Marymount Hospital, 12300 McCracken Road, Garfield Heights, Cleveland, 44125, OH, USA
[c] Center for Bioethics, Clinical Transformation, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH, USA
[d] Office of Nursing Research and Innovation, Stanley S. Zielony Institute for Nursing Excellence, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH, USA

ARTICLE INFO

Article history:
Received 13 November 2020
Received in revised form 13 May 2021
Accepted 14 May 2021

Keywords:
Moral distress
Nursing
Clinical ethics
Bioethics
Systematic review

ABSTRACT

Background: Moral distress is a pervasive phenomenon that can negatively impact healthcare professionals and has been well studied in nursing populations. Much of the evidence suggests that it is associated with intention to leave high acuity areas and the profession. Despite the increasing amounts of research to explore the causes and effects of moral distress, there is limited research on interventions that mitigate the negative effects of moral distress.
Objectives: The aims of this systematic review were to: (a) identify and examine interventions developed to address moral distress experienced by health care professionals (b) examine the quality of the research methods and (c) report on the efficacy of these interventions.
Design: We conducted a systematic review of interventional studies developed to mitigate moral distress.
Data Source: Medline, Embase, PsycINFO, CINAHL and Cochrane were searched for relevant studies (July 2019- September 2019). Additional bioethics databases and reference lists were also hand-searched.
Review methods: The first author reviewed all retrieved titles and abstracts with a low tolerance for borderline papers based on inclusion and exclusion criteria, and those papers were reviewed and discussed by all authors to determine inclusion. Quality appraisal was conducted on the included studies using narrative synthesis to compare the findings. Data were extracted and compared by all authors and then reviewed by the first author for consistency.
Results: Sixteen papers were included for full text review and the following interventions identified: educational interventions of varying length and breadth; facilitated discussions ranging from 30 to 60 minutes; specialist consultation services; an intervention bundle; multidisciplinary rounds; self-reflection and narrative writing. Researchers reported statistically significant reductions in moral distress using pre and post surveys, including one mixed methods program evaluation *(n=7).* The qualitative program evaluation provided participant quotations to suggest their program was beneficial. There were no statistically significant findings in the other studies *(n=8).* All studies had limitations in design and methodology presenting significant threats to validity.
Conclusion: Designing rigorous research studies that measure the impact of interventions aimed at mitigating moral distress continues to be challenging. The primary reason being that moral distress is a subjective ethical phenomenon with a number of different causes and effects. This calls for interventions that are flexible and sensitive to individual's needs. To build an evidence-base, interventions should also be measurable and research methods need to be scientifically rigorous. To achieve rigor and innovation, researchers should clearly justify their methodological choices.
Tweetable abstract: Interventions to mitigate moral distress: a systematic review of the literature. Educational interventions offer a promising direction but more research is needed.

* Corresponding author.
E-mail addresses: morleyg@ccf.org (G. Morley), fieldr@ccf.org (R. Field), colec@ccf.org (C.C. Horsburgh), Christian.Burchill@pennmedicine.upenn.edu (C. Burchill).
** Currently Director of Nursing Research and Science, Penn Medicine Lancaster General Hospital, Lancaster, Pennsylvania, United States of America.

https://doi.org/10.1016/j.ijnurstu.2021.103984
0020-7489/

What is already known

- Moral distress negatively impacts nurses and other healthcare professionals both personally and professionally.
- Whilst there is an abundance of research that evidences the impact of moral distress, there is a dearth of literature that reports evidence-based, effective interventions that mitigate the negative effects of moral distress.

What this paper adds

- There are a number of interventions developed with the specific aim of mitigating moral distress; however, quality of interventions and research designs are highly variable and often lack rigor.
- Single group designs that lack a comparator group are frequently used in studies of moral distress interventions with researchers claiming that results demonstrate significant or insignificant effect of the intervention.
- Educational interventions formed the most common intervention tested. Although intervention fidelity was low and bias high, many researchers found a statistically significant reduction in moral distress that might suggest this as a promising direction.

1. Introduction

Moral distress was first highlighted in the nursing literature by Jameton (1984) who stated that it occurs when, "one knows the right thing to do, but institutional constraints make it nearly impossible to pursue the right action" (p.6). Over the past three decades, moral distress has received increasing amounts of research attention. More recently, due to the lack of clarity in empirical papers, philosophers and researchers have begun to debate the meaning of the concept itself. Some of these authors find Jameton's definition to be too narrow and argue for a broader understanding based upon conceptual arguments (Fourie, 2015; Campbell et al., 2016), empirical data (Morley et al., 2020) or both (Morley, Bradbury-Jones, & Ives, 2021). At present there is no singular agreed upon definition. There is some recognition that moral distress may have some positive value because it shows healthcare professionals are emotionally engaged with ethical issues (Tigard, 2018). Nonetheless, there is broad agreement that the negative effects of moral distress are problematic both personally and professionally since it causes healthcare professionals to temper their emotional connection to patients and families, and to consider leaving high acuity areas and their profession (Rushton et al., 2015; Helft et al., 2009; Morley, 2018). There is also evidence to suggest moral distress may be correlated to other negative concepts such as burnout and compassion fatigue (Maiden et al., 2011; Rushton et al., 2015; Neumann et al., 2018).

Based on Jameton's constraint-based definition, researchers began to explore the concept of moral distress in populations that they hypothesized would experience it most frequently and intensely such as intensive care unit clinicians and, more recently, amongst palliative care clinicians (McAndrew et al., 2018; Maffoni et al., 2019). Since nurses frequently experience constraints on their ability to make decisions because of power structures within healthcare systems, moral distress research has occurred primarily amongst the nursing population but this has now expanded to cohorts of other healthcare professionals (Allen et al., 2013; Whitehead et al., 2014). The moral distress field is now moving from descriptive-correlational designs about prevalence, severity, and effects of moral distress to developing and testing interventions that mitigate the negative effects on healthcare professionals. Burston and Tuckett (2012) reported on a number of publications with suggested interventions to address moral distress but implementation of interventions and research to evidence effects remain largely in development. Nonetheless, there are enough published studies to warrant a systematic review of the literature to date. We use the language of mitigating and/or addressing moral distress because we recognize that moral distress may be a regarded as a normal response to morally troubling events and is a consequence of a pluralistic society in which individuals prioritize values differently, and therefore eradicating moral distress is likely impossible and undesirable (Nyholm, 2016; Howe, 2017; Tigard, 2018). We suggest the aim ought to be mitigation or reduction of the *negative effects* of moral distress although we recognize that some of the studies we reviewed aimed to reduce moral distress conceived of as a whole phenomenon.

Aims of this systematic review were to (a) identify and examine interventions developed to address moral distress experienced by healthcare professionals, (b) examine the quality of the research methods, (c) and report on the efficacy of these interventions. There is no protocol associated with this review.

2. Methods

2.1. Search strategy

We used a systematic approach to the review following the 7-step approach suggested by Strech at al. (2008) and integrated guidance from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2009), Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2019), and the Template for Intervention Description and Replication Checklist (Hoffman et al., 2014).

Strech et al. (2008) suggested a 7-step approach to formulating systematic reviews of empirical bioethics literature: (1) careful definition of the review question; (2) selection of relevant databases; (3) application of ancillary search strategies; (4) development of search algorithms; (5) relevance assessment of the retrieved references; (6) quality assessment of included studies; and (7) data analysis and presentation. Although the authors of the included papers were not making explicit normative recommendations, we anticipated the need to integrate ethical concepts and empirical methods in the review, so Strech et al.'s approach provided a robust strategy for considering these elements. However, once we completed initial data extraction we found that very few researchers engaged with the ethical concepts. Nonetheless, with integration from PRISMA (Moher et al., 2009), Cochrane (Higgins et al., 2019) and TIDieR (Hoffman et al., 2014), the process we followed was sufficiently systematic and transparent. Strech et al. (2008) suggested first formulating a review question. Our review question was, 'What interventions have been developed to address, mitigate, or reduce moral distress experienced by healthcare professionals?'

According to Strech et al.'s methodology, the review question also informs steps 2, 3 and 4: the search of relevant databases, application of ancillary search strategies and the search algorithm. The search strategy with Boolean operators used in Medline is shown in Table 1. The same keywords were used but with MeSH terms adjusted according to the particular database. We ran this search in Ovid Medline, Embase, PsycINFO, CINAHL, and Cochrane Library. We limited our search results from 1999 onwards to make the number of results manageable for screening. We felt confident there would be few intervention studies that would pre-date 1999 since the concept was first mentioned in 1984. The search was conducted from July 2019- September 2019. We also handsearched reference lists for relevant papers with the most recent hand search in January 2021. We also hand searched bioethics databases that do not have advanced search functions such as

Table 1
Search strategy in medline.

1 moral* distress*
2 (moral* adj (dilemma* or challeng* or injur* or uncertaint* or conflict* or constraint*))
3 morals/
4 exp stress, psychological/
5 3 and 4
6 exp Health Personnel/
7 exp Students, Health Occupations/
8 exp religious personnel/ or social workers/
9 ((health or healthcare or "health care" or hospital) adj (professional* or practitioner* or personnel))
10 (physician$1 or doctor$1 or clinician$1 or provider$1)
11 (nurse or nurses or nursing)
12 (medicine or medical or pharmacist$1 or "social worker$1")
13 (chaplain$1 or clergy or "spiritual care")
14 (1 or 2 or 5) and (or/6-13)
15 (manag* or intervention* or mitigat* or reduc* or therap*)
16 (treat* or cope or coping or decreas* or strateg* or training or tool or tools)
17 14 and therapy.
18 14 and (15 or 16)
19 17 or 18
20 limit 19 to (english language and yr="1999 -Current")

the Bioethics Research Library of the Kennedy Institute of Ethics, EthxWeb, and EUROETHICS.

2.2. Inclusion and exclusion criteria

Previous research has shown a lack of consistency between the way in which moral distress is defined and the items used in the measure of moral distress (Hamric, 2012). For the purposes of this review, moral distress could be defined by the researchers who conducted the study. However, we also examined whether the definition was consistent with both the measure of moral distress and the intervention to ensure it reflected the same underlying concept and was consistent. Our inclusion criteria were intentionally broad and included studies using different interventions and outcomes measures due to the limited number of published research studies examining interventions addressing moral distress. Inclusion criteria were: (a) use of an intervention implemented to mitigate moral distress in healthcare workers; (b) at minimum, post-intervention evaluation of moral distress; (c) and text available in English. Studies were intentionally not limited to randomized controlled trials or quantitative methods due to the limited number of studies. Quality improvement, feasibility, pilot, program evaluation, and qualitative studies were included. We wanted to be open to the inclusion of interventions aimed at all healthcare professionals, but many of the included studies were with nursing populations and so this became a focus of the review. There were no exclusion criteria.

2.3. Study selection

After removing duplicates, all titles and abstracts were reviewed by the first author based on the inclusion and exclusion criteria. Although we recognize that review by two individuals would have been preferable, the first author had a low tolerance for borderline studies that would then be reviewed by the research team to assess suitability. Any disagreement was resolved within the team through consensus.

2.4. Data extraction

The first author developed a data extraction table. The data extraction table was based on the PRISMA checklist (Moher et al., 2009), TIDieR checklist (Hoffmann et al., 2014) and the aims of the review, which were to report on the intervention, the extent to which moral distress was addressed and the quality of the research project. The papers were initially divided between the research team and each reviewer was responsible for reporting on their designated studies. The research group came together to discuss the strengths and weaknesses of their initially designated papers and reviewed the remaining papers.

2.5. Quality appraisal and risk of bias

Quality appraisal is another area susceptible to bias and many checklists exist to mitigate this bias (Strech et al., 2008). Each study was appraised based upon criteria adapted from the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2019), TIDieR (Hoffman et al., 2014), and PRISMA (Moher et al., 2009). We integrated these reporting guidelines because we aimed to follow a systematic approach for the review itself, and to provide sufficient analysis of the interventions themselves.

Key quality indicators were: (a) risk of bias; (b) sufficient explanation of the intervention; (c) replicability of the intervention; (d) consistency between definition of moral distress, intervention and measure; (e) validity of measure to assess moral distress; (f) data analysis; (g) report of findings.

2.6. Data synthesis

A narrative summary technique was used to describe the review findings rather than a meta-analysis because of the variability of the reporting on statistical findings and between interventions (approach, duration, and length). A narrative approach was also deemed the most suitable because a key aim of the review was to summarize and describe each intervention and a meta-analysis would have reduced the richness of our report. To synthesize the data, we followed the following steps suggested by Popay et al. (2006): (1) developed a preliminary synthesis of the studies and the effects; (2) explored relationships between the studies, similarities and differences and we considered the strengths and limitations of various research designs and interventions; (3) assessed the robustness of the included papers and the synthesis itself. A fourth optional step suggested by Popay et al. (2006) is to develop a theory of change in which the theory behind the studies is described. We did not find one unifying theory because many of the authors adopted different underpinning theories for their interven-

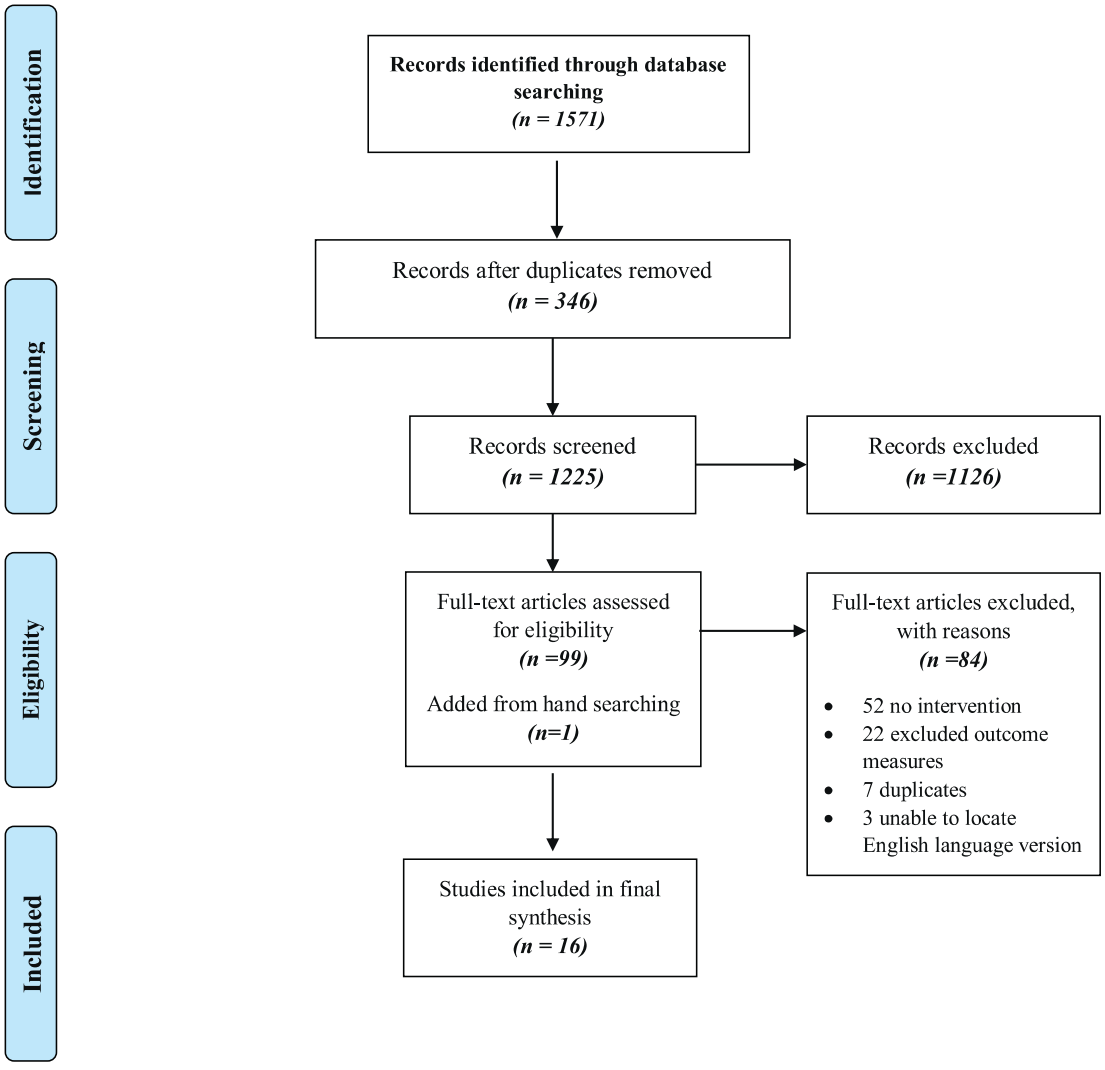

Fig. 1. PRISMA flow diagram.

tions. We report on these theoretical underpinnings and pay close attention to consistency.

3. Results

3.1. Study selection

A total of 1571 records were identified from the electronic databases and, after removing duplicates, 1225 abstracts were screened and 16 papers included for full text review. Due to the first author's familiarity with the moral distress literature, this process was conducted by one author. Fig. 1 shows the PRISMA flow diagram. Study designs that would usually be excluded from a robust systematic review were included – notably two program evaluations, pilot study and feasibility trial – because we aimed to describe interventions and author's inferred changes in moral distress. We included four pilot studies (Browning & Cruz, 2018; Fontenot & White, 2019; Meziane et al., 2018; Wocial et al., 2017), two program evaluations (Hamric & Epstein, 2017; Robinson et al., 2014), one quality improvement initiative (Vaclavik et al., 2018), three single-group design research studies (Brandon et al., 2014; Monteverde, 2016; Rushton et al., 2021), two mixed methods studies (Allen & Butler, 2016; Chiafery et al., 2018), one quasi-experimental study (Beumer, 2008), and three randomized trials (Abassi et al., 2018; Molazem et al., 2013; Saaedi et al., 2018). Table 2 presents data extraction and appraisal results.

3.2. Risk of bias

Risk of bias was high. Most commonly, authors reported a single group research design using a convenience sample in which researchers appeared to be interventionists *(n = 11)*. Of authors who reported randomized controlled trials (Abassi et al., 2018; Mozalam et al., 2013; Saaedi et al., 2018), Saeedi et al. (2018) reported single blinding, loss to follow up, and intent-to-treat analysis. No studies reported a priori sample size calculations or post hoc power analysis. Table 3 presents risk of bias analysis.

3.3. Measures of Moral Distress

The Moral Distress Scale (MDS) (Corley, 1995) and Moral Distress Scale-Revised (MDS-R) (Hamric et al., 2012) were the most frequently used outcome measures followed by the Moral Distress Thermometer (MDT) (Wocial & Weaver, 2013). Three studies reported using previously validated Persian language versions of the MDS or MDS-R (Abassi et al., 2018; Molazem et al. 2013;

Table 2
Data extraction table.

Author (year), Location	Study Design	Participants	Outcome Measure	Intervention	Results	Appraisal
Abbasi et al. (2019), Iran	Randomized clinical trial (randomization at the individual participant level)	ICU nurses *(n=60)* from a teaching hospital I/C: 30/30	Moral Distress Scale-Revised (MDS-R), translated into Persian with validity and reliability reported. Cronbach's alpha 0.84 for frequency of MD; 0.82 for intensity; 0.86 for the total score of scale in Persian version.	**Educational intervention**: 2 sessions x 6 hours for intervention group.1 × 2-hour session for control that did not include content about strategies to manage MD.	Reported a significant difference in the intervention group *(p<0.05)*, with the least significant difference between pre- and 2-weeks post, and the most significant difference between pre-and 4-weeks post.	No sample size calculation a priori or post hoc, no report of response rate to surveys. Although control received education, it is not comparable (2 hours compared to 12). MDS-R not designed to test unique moments of MD but rather to assess more global or 'chronic' MD so unclear whether this is the most effective measure to use.
Allen & Butler (2016), USA	Mixed methods cross-sectional survey design with one group pre-and post-intervention testing, and focus group interviews.	Convenience sample of adult critical care nurses *(n=12)* and pediatric nurses *(n=7)*, from a tertiary care hospital.	MDS-R for adult & pediatric samples pre and 3-months post; Hospital Ethical Climate Survey (Olson, 1998) pre-intervention; focus group pre and 3-months post.	**Educational intervention**: '2-hour blended learning training' developed from focus groups with nurses *(n=4)*.	*n=4* completed the MDS-R 3-months post intervention and *n=1* reported a reduction in MD using the MDS-R at 3-month follow up. Focus group participants reported that they found the education helpful.	Small sample size, drawn from only one unit in one hospital. The authors state the findings aren't generalizable however the aim of a pilot is not to create generalizable knowledge but to assess feasibility. The authors do not address feasibility. The education session was developed from focus groups and may be effective as derived from their experiences.
Beumer (2008), USA	Quasi-experimental design	Convenience sample of regular staff for intervention & supplemental staff for control, single setting, ICU nurses *(n=34)*. Hospital type not stated. I/C: 21/13	Moral Distress Questionnaire developed by author with 8 Likert scale questions and 4 true/false statements. Validity or reliability was not tested; psychometrics not reported.	**Educational intervention**: designed & presented by nurse manager, employee assistance counselor and clinical nurse specialist. 2 hour workshop. No intervention for control group.	Post questionnaire completed at 7-10. No statistical analysis of change completed. Both intervention & control groups reported % increased responses that nurses' opinions are valued & they have resources to cope with morally distressing situations. Decrease in distancing from patients in intervention group.	No validity or reliability testing of questionnaire, no psychometrics reported. Not clear that the questionnaire captures MD. Intervention & control group not comparable.
Brandon et al. (2014), USA	Pre/ post intervention cross-sectional web-based survey	Convenience sample of pediatric in-patient and out-patient HCP (pre: *n=413);* (post: *n=364).*Academic tertiary medical center.	Modified version of Corley's Moral Distress Scale (2001). Reliability and validity measurement not reported.	**Specialist consultation bundle:** Pediatric Quality of Life (QoL) program implemented: including specialist consultation, care coordination, symptom management, end-of-life care, patient education, spiritual car & debriefing sessions for difficult cases.	Scale administered pre & 20 months post. Small reduction in mean MD scores across all items- two sub themes when adjusted for role had statistically significant reductions *(p<0.05):* frequency of which participants felt they were providing care not in the patient's best interests & reduced intensity of MD related to work QoL.	Did not pair testing pre & post. Authors altered a valid & reliable scale without repeated testing. No psychometrics reported on work QoL survey. Respondents only had to answer 50% of scale items.

(*Continued on next page*)

Table 2 (*Continued*).

Author (year), Location	Study Design	Participants	Outcome Measure	Intervention	Results	Appraisal
Browning and Cruz (2018), USA	Pilot study. Pre/post-test experimental design	Convenience sample of ICU nurses *(n=42)* of which *(n=6)* participated in the intervention & completed pre & post surveys. Type of hospital not specified. I/C: 19 /23	MDS-R with added question about participant preferences to continue debriefs.	**Facilitated discussion:** 'Reflective Debriefing' facilitated by a social worker with an ethics education component. Case chosen by nurse manager and nurses on the unit based on the amount of 'stress' caused. Sessions lasted 45-60 mins with 5-10 participants. Control group did not participate in the intervention.	Pre-and post, nurses reported low to moderate levels of overall MD & same top scoring frequency items. No statistically significant differences in MD levels in intervention group. Participants reported finding sessions helpful & 63% wanted them monthly.	Stated method is pilot study but no mention of feasibility. Very small sample size. One group design for pre-and-post testing. No paired testing pre-post. Due to the educational component, unclear if testing discussion or education. Intervention effectiveness likely to be dependent on the skillfulness of the facilitator.
Chiafery et al. (2018), USA	Pre/post intervention study, one group mixed methods design.	Convenience sample of adult critical care nurses *(n=32)* from 3 ICUs, tertiary academic medical center.	Moral Distress Thermometer (MDT) at 3 time points (baseline, pre & post); web-based free text survey one week post; interviews with participant managers.	**Facilitated discussion:** Nursing Ethics Huddles (NEH): approximately 30 minutes in length. Nurse ethicist facilities group discussion of a patient case that includes reflection, clarification of stakeholder values & ethics education.	68% reported reduced MD using MDT, 18% reported increase, 14% unchanged. Statistically significant difference in MDT pre- and post-scores *($t=3.55$, $p< 0.01$, Cohen's $d=0.7$)* indicating moderate effect. Free text responses were favorable, 50% respondents stated they would participate again.	Small sample size. Risk to consistency because it is not entirely clear which definition of MD was used with the MDT. Figure 2 presenting the process of NEH is unclear.
Hamric & Epstein (2017), USA	Program evaluation.	Purposive sample from consult requestors between 2006-2012 *(n=19);* 2013-2016 *(n=40)* from in-patient and out-patient settings in a tertiary academic center.	3-month post consultation interviews with requestors (74% response rate). MDT used pre- and post-consultation from 2015-2016.	**Specialist consultation:** Moral Distress Consultation Service developed to address MD separately from the Ethics Consultation Service. Facilitation by ethics team members for 1 hour sessions.	83% of interview respondents indicated the consult led to resolution of key problems, changes in staff or team behavior, or improved organizational processes.	Insufficient report of qualitative research methods. Training required to conduct intervention not explained.
Fontenot & White (2019), USA	Pilot study with pre/post one group design.	Purposive sample of critical care nurses *(n=21)* from one unit, large academic medical center.	MDT pre and post.	**Facilitated discussion**: 'Evidence-based debriefing'. Sessions lasted 30 minutes, offered every 2 weeks for 10 weeks. Facilitated by a social worker with training in group therapy and MD.	*(n=13)* completed pre-and post- surveys and unclear how many debriefs each participant attended. Paired t-tests found no significant difference between pre- and post MD scores. MD incrementally rose from session1-3 and then back down to pre-debrief scores.	Authors could not use linear modelling for statistical analysis. Authors provide a brief content guide but do not provide facilitation skills required. Educational component mentioned but not explained. Pilot method stated but do not address feasibility.
Meziane et al. (2018), Canada	Pilot study with one group design.	Convenience sample of nurses *(n-19)* from acute care units of a university hospital.	MDS-R pre and 2-weeks post.	**Self-refection:** Reflective practice (RP) intervention. 3 sessions provided lasting 45-75 mins. Session facilitated by a palliative care clinical nurse specialist.	No statistically significant effect on MD. 2 participants dropped out due to availability, 4 lacked the time to complete the reflective journaling. 4 participants reported that journaling was tedious. Study is feasible given recruitment & attrition but no evidence of effect.	One group design makes it difficult to determine effect of intervention. Authors acknowledge that the MDS-R was not designed for use as a pre- and post-design measure.

(*Continued on next page*)

Table 2 (*Continued*).

Author (year), Location	Study Design	Participants	Outcome Measure	Intervention	Results	Appraisal
Molazem et al. (2013), Iran	Block randomized controlled trial.	Convenience sample of cardiac critical care nurses (*n=60)* from one hospital. I/C: 30 /30	Moral Distress Scale, translated into Persian. Authors state it was tested for reliability & validity with Iranian nursing population in a previous study.	**Educational intervention**: 2 four-hour sessions. Authors did not disclose expertise of educators and facilitators. No intervention for control group.	Completed pre, 1 and 2-months post. Significant difference between mean MD scores at 1 & 2 months. Control group: increased, 4.712±1.048 to 5.275±0.946 (1 month) and 5.183±1.153 (2 months). Intervention group: decreased, 4.441.24 to 3.36±0.996 (1 month) and 3.048±1.25 (2 months).	Authors are relying on participants not sharing education with the control group as they used one unit for both groups. The sample may not be representative: from one unit and one hospital. The authors do not report psychometrics for the MDS.
Monteverde (2016), Switzerland	Explorative, quantitative pre/post interventional one-group design.	Purposive sample of nursing students *(n=166)* from 3 baccalaureate programs, one university.	MDT translated into German and pre-tested, reliability validity not reported. Moral resilience hypothesized as absence of MD.	**Educational intervention**: 30 minute lecture aimed at differentiating between morally complex and morally wrong situations.	MDT pre & post. In 3 of 4 vignettes, there was a modest, statistically significant reduction in MD ($p< 0.05$, $\propto = 5\%$). Wilcoxon two-tailed ranks test was used $\propto = 5\%$ for statistical analysis.	Vignettes developed from former student's narratives so should be relatable. Risk to internal validity because the same vignettes were used pre-and post but utilizing different vignettes would reduce comparability.
Robinson et al. (2014) USA	Program evaluation with pre/post interventional one-group design.	Three cohorts of registered nurses *(n=67)* from 2 academic medical centers over 3 year period in 3 cohorts: clinical nurses *(n=46)*, advanced practice nurses *(n=11)*, nurse leaders *(n=10)* with the latter in separate groups.	MDS-R, Ethics Knowledge Scale & Self-Efficacy Scale in Clinical Ethics. Latter two author developed. Free text narratives also invited.	**Educational intervention**: 2-hour virtual foundational course, 80-hours of classroom teaching, 16-hours of additional clinical mentorship.	Pre & 2-weeks post last classroom day. Paired t-tests conducted to evaluate impact on MD. Statistically significant reduction in MD from time 1 [M= 72.04; SD=33.59] to time 2 [M= 56.82; SD= 29.29] *(p<.000)*. Increased knowledge *(p<.005)* and self-efficacy *(p<.000)*. Qualitative findings supported the value of the program.	Sample was divided into three cohorts over three years & results reported as one thereby reducing ability to see variation between cohorts. Authors present limited MD data, presenting on the paired t tests, mean, standard deviation and p values. The mixed methods approach could have been enriched by highlighting how the qualitative reports mirrored/ were correlated to the items in the MDS-R.
Rushton, et al. (2021), USA	Pre/post interventional one-group design.	Convenience sample of nurses form different clinical areas from 2 hospitals in a large academic medical system, recruited from 2016-2018 in 2 cohorts. I/C: 192/223.	MDT, Perceived Ethical Confidence Scale, Moral Sensitivity Questionnaire, Moral Competence Questionnaire, Brief Resilience Scale, Multidimensional Emotional Empathy Scale, Work Engagement, 2-item burnout questions, 1-item turnover intention (modified), Ilfeld Psychiatric Symptom Index, Mindful Attention Awareness Scale.	**Education intervention:** 6 experiential sessions totaling 24 hours of face-to-face education and training; 1 session involved high-fidelity simulation with trained actors and a facilitated reflective debrief. No intervention for control group.	Independent t tests and x squared tests to determine differences between 2 hospitals & evaluated impact with repeated analysis of covariance. No difference in MD. Significant increases in ethical confidence, ethical competence, resilience, work engagement, mindful awareness & attention.	Sample was divided into two cohorts over two years & results reported as one thereby reducing ability to see variation between cohorts. Participants received eleven survey measures increasing possibility of response fatigue.

(*Continued on next page*)

Table 2 (*Continued*).

Author (year), Location	Study Design	Participants	Outcome Measure	Intervention	Results	Appraisal
Saeedi et al. (2018), Iran	Block randomized controlled trial.	Purposive sampling of ICU and NICU nurses *(n=106)* from 2 teaching hospitals. I/C: 55/51.	MDS translated into Persian, content validity index of 88% and Cronbach's alpha coefficient of 90%.	**Narrative writing:** Once a week (at minimum) over an 8-week period. No intervention for control group.	There were no statistically significant findings in terms of reduction in MD, frequency or intensity.	Threat t consistency, authors defined 'moral stress'. Lack of adequate time to engage in intervention due to workload & lack of protected time. Insufficient information provided on narrative writing education.
Vaclavik et al. (2018), USA.	Quality improvement with pre/ post intervention evaluation, one group design.	Convenience sample of adult inpatient hematology oncology nurses *(n=56)* from an academic medical center. Pre-intervention group *(n=28)* and post intervention group *(n=18)*.	MDS-R.	**Intervention bundle:** 'bundle of mindfulness interventions' based upon mindfulness-based stress reduction (MBSR) approaches.	Pre-intervention & 3-months post. Reduction of MD related to one item only: 'how frequently does the distressing situation (witnessing healthcare providers giving false hope to a patient or family) occur'. They state frequency of MD due to this item reduced from 81% to 44%.	Interventions implemented at the same time so unable to tell if certain interventions are more impactful than others. Do not report all findings but cherry pick two items to report with no clear rationale. State in discussion that the critical debriefs & work-life balance events had the greatest impact but there is no evidence provided to support this claim.
Wocial et al. (2017) USA.	Pre-and post-pilot intervention, one group design.	Convenience sample of PICU multi-disciplinary staff from a quaternary care hospital. Pre-intervention group *(n=131)*; post intervention group *(n=89)*; pre and post *(n=42)*.	MDT, MDS-R, PEACE Discussion Evaluation Form, latter author developed.	**Multidisciplinary rounds:** boardroom rounds to discuss ethical issues & communication coaching provided. Once a week over a 12 month period.	MDT scores fluctuated throughout. MDS-R overall scores were lower for respondents in all categories (non-significant) and on three specific items (significant). Greater effect on nurse MD than physicians.	Unable to compare pre-and post MDS-R scores with attendance. Limited data reported from MDT scores bi-monthly. No longitudinal follow up. Discussion & facilitation dependent on a skilled clinical ethicist able to engage all members of the team. Cost effective intervention although does require buy-in & time investment.

Saeedi et al., 2018). Saeedi et al. (2018) was the only author group to report adequate instrument validity testing results in their study. Brandon et al. (2014) reported altering Corley's original MDS retaining four original items without reporting validity and reliability of this version. Wocial et al. (2018) used both the MDS-R, pre-/post-intervention, and MDT throughout the intervention. The MDT was used exclusively in four studies (Chiafery et al., 2018; Fontenot & White, 2019; Monteverde, 2016; Rushton et al., 2021). Beumer (2008) developed a questionnaire that consisted of five Likert-like scale items and four true/false items based upon a review of the moral distress literature but did not report psychometric results. Hamric & Epstein (2017) conducted interviews 3-months post-consultation but did not report data collection or analysis methods.

3.4. Intervention characteristics

3.4.1. Facilitated discussion

Facilitated discussion was implemented as an intervention to decrease moral distress in three studies. Facilitated discussions were often based on theoretical or educational models and included educational instructions and discussion about recent morally distressing events. Three different theoretical approaches were used to frame these interventions. Browning & Cruz (2018) based their facilitated discussion intervention, Reflective Debriefing, on a review of moral distress literature as well as the educational debriefing literature, integrating components of both fields into one intervention. Chiafery et al. (2018) used Nathaniel's *Theory of Moral Reckoning* and Rhodes and Alfandre's (Rhodes and Alfandre, 2007) clinical ethics model for their intervention. Fontenot & White (2019) based their intervention on American Association of Critical-Care Nurses' *4As to Rise above Moral Distress* model (American Association of Critical-Care Nurses 2004). Facilitated discussion interventions often involved recent incidents in which unit leaders thought moral distress might have occurred and an unstructured discussion of the case was developed in real time by the person leading the intervention so that debriefing and learning could occur. Sessions lasted between 30-60 minutes (Browning & Cruz, 2018; Chiafery et al., 2018; Fontenot & White, 2019). Interventions were delivered once a month for six months (Browning & Cruz, 2018), at unspecified intervals over two months (Chiafery et al., 2018), or every two weeks for 10 weeks

Table 3
Risk of Bias.

Reference and location	Researcher bias	Selection bias	Attention bias	Reporting bias	Selective reporting	Threat to Validity
Abbasi et al. (2018), Iran	-	-	-	-	+	-
Allen & Butler (2016), USA	?	-	+	+	-	+
Beumer (2008), USA	?	-	+	+	+	+
Brandon et al. (2014), USA	?	-	-	-	+	+
Browning and Cruz (2018), USA	?	+	+	-	-	+
Chiafery et al. (2018), USA	+	+	N/A	-	+	-
Hamric & Epstein (2017), USA	+	-	N/A	-	+	+
Fontenot & White (2019), USA	-	+	N/A	-	-	+
Meziane et al. (2018), Canada	?	+	N/A	-	-	-
Molazem et al. (2013), Iran	-	-	+	+	-	+
Monteverde (2016), Switzerland	?	-	N/A	+	-	-
Robinson et al. (2014) USA	?	+	N/A	-	+	-
Rushton et al. (2021)	+	+	+	-	-	-
Saeedi et al. (2018), Iran	-	-	+	+	-	+
Vaclavik et al. (2018), USA.	?	+	N/A	+	+	+
Wocial et al. (2017), USA.	?	-	N/A	-	-	-

Key: - low risk of bias; + high risk of bias; ? unclear risk of bias; N/A not applicable in study context.

(Fontenot & White, 2019). In two studies it appeared that the intervention was delivered by the primary investigators, a clinical social worker experienced in group facilitation (Browning & Cruz, 2018) and a nurse ethicist (Chiafery et al., 2018).

3.4.2. Self-reflection

Meziane et al. (2018) described a reflective practice intervention to help nurses come to terms with contradictions that exist between the ideal and actual care provided based on Johns' (Johns, 2006) model for structured reflection and Watson's concept of human caring. The intervention occurred in three phases: phase one included in-person education on identifying moral distress and Johns' (Johns, 2006) model for reflective practice followed by participants writing a structured self-reflection; phase two involved instruction on performing reflective practices using the written self-reflection; and phase three involved lessons on applying strategies to deal with moral distress. All sessions were investigator-led every two to three weeks lasting 45-75 minutes.

3.4.3. Narrative writing

Saeedi et al. (2018) used a narrative writing intervention to address moral distress. Authors provided few details on intervention content and teaching method except that participants were provided with a definition of moral distress, examples of narrative writing, a notebook, and instructions to write about situations that caused moral distress. Authors reported that notebooks were checked by a researcher to ensure participants were following instruction, but did not provide detail about how the overall process was used to address moral distress.

3.4.4. Intervention bundle

Vaclavik et al. (2018) reported on a bundle of interventions designed to increase resilience and reduce moral distress. The bundle included: critical debriefing led by a licensed grief counsellor following a critical event; a bag containing a lavender sachet, chocolate bar, and tissues; a *tree of life* wall art installation in the unit break room on which names of patients who died were displayed; establishment of a work-life balance committee; yoga classes offered on the unit on both shifts twice a week; and training in mindfulness-based stress reduction techniques. Of note, the only intervention in the bundle substantiated with supportive literature was the mindfulness-based stress reduction class.

3.4.5. Multidisciplinary rounds

Wocial et al. (2017) utilized a semi-structured interprofessional-rounding format intended to promote open discussion about ethically challenging situations and moral distress. The meeting, named Pediatric Ethics and Communication Excellence (PEACE) Rounds, was a process-oriented format for clarifying decision making and achieving consensus around broad goals of treatment for children with life-threatening illness in the pediatric intensive care unit. Rounds consisted of a synopsis of the care plan for a case in which the healthcare team felt that there may be incongruence between family and team member expectations followed by a discussion facilitated by an ethicist who helped negotiate disagreements on care when they arose and offered education and coaching on communication techniques used for difficult conversations.

3.4.6. Specialist consultation service programs

Two studies evaluated consultation programs as a way to address moral distress. Brandon et al. (2014) described a newly developed program that included consultation services from two physicians and one pediatric nurse practitioner in palliative care. Consultations were available to clinicians throughout the healthcare system for patient care planning and coordination needs as well as educational offerings, debriefings and weekly unit-based interdisciplinary meetings to discuss specific patient care issues in real time. Hamric and Epstein (2017) developed a moral distress consultation service staffed by members of their ethics consultation service, though education and training background of consultants is absent from the report. Authors described the consultation process as a group process in which one facilitator established boundaries for a safe environment in which attendees shared thoughts and feelings and the other acted as scribe. Cases were discussed by attendees with the facilitator clarifying value differences and constraints so that an action plan would be put in place that participants can agree is 'right' based upon professional values and standards.

3.4.7. Education interventions

Authors who developed interactive educational workshops employed a variety of frameworks to underpin their interventions. Investigators often also served as the educators. Two authors used the American Association of Critical-Care Nurses' (American Association of Critical-Care Nurses 2004) *4A's to Rise above Moral Distress* model for their intervention. Beumer (2008) provided five workshops over a 4-week period totaling 10 hours during which participants received instruction on how to recognize, cope with and address moral distress. Molazem et al. (2013) provided two 4-hour educational workshops delivered over two consecutive weeks with similar content: defining and discussing moral distress; developing individual and system strategies to address it; applying the *4 A's* framework, and role-playing and practicing responses to moral distress scenarios.

Other studies reported different educational models. Monteverde (2016) provided education on the concepts of moral complexity and moral wrongness during a 30-minute lecture followed by a short discussion during which participants were encouraged to provide examples from practice. Allen & Butler (2016) reported using a blended learning training model that involved focus group discussion but failed to provide details of group discussion guidelines and educational curriculum. Abbasi et al. (2019) tested a 2-day workshop based on Nathaniel's *Theory of Moral Reckoning in Nursing*. Authors provide detail of the training techniques and content that match the stages in the theory.

Robinson et al. (2014) and Rushton et al. (2021) developed the most robust educational interventions. Robinson et al.'s (2014) Clinical Ethics Residency for Nurses (CERN) consisted of 98-hours of education and mentorship over a ten month period. The curriculum was developed from moral philosophy, US professional standards for bioethics and ethics consultation (drawn from the American Society for Bioethics and Humanities (2009) and the American Nurses Association Code of Ethics (American Nurses Association, 2015). The curriculum for clinical nurses was aimed at increasing their ability to detect emerging ethical issues, initiate preventative actions, lead ethics rounds and serve as unit ethics resources. The education consisted of a 2-hour virtual foundational course, 80-hours of classroom teaching that included exploration and sharing of values, communication techniques, role-play, simulation, and mentorship. By contrast, Rushton et al.'s (2021) Mindful Ethical Practice and Resilience Academy (MEPRA) totaled 24-hours face-to-face and consisted of didactics, interactive training, role-play, facilitated reflective debriefing and mindfulness practices. The focus of CERN seemed to be upon providing in-depth ethics education whereas MEPRA focused more upon increasing ethical confidence and teaching individualized coping mechanisms.

3.5. Intervention effects

3.5.1. Pilot and feasibility studies

Authors of pilot studies in which changes in moral distress pre-to-post-intervention were tested found no significant difference following the intervention (Browning & Cruz, 2018; Fontenot & White, 2019; Meziane et al., 2018) and only Meziane et al. (2018) fully reported feasibility results.

3.5.2. Program evaluation

Hamric and Epstein (2017) conducted post-intervention interviews reporting that the vast majority of participants felt that the moral distress consultation led to resolution of the events that created a need for a moral distress consult. Themes identified by authors that supported their conclusion that the program was successful included: acknowledgement of staff concerns; staff empowerment; staff engagement; improved team collaboration, and unit or organizational-level change. Robinson et al. (2014) used both quantitative and qualitative methods to support utility of their program. They found statistically significant pre- to post-intervention in decreased MDS-R total score, increased knowledge, and increased self-efficacy. They reported qualitative results from participants that supported quantitative findings.

3.5.3. Quality improvement

Vaclavik et al. (2018) reported a statistically significant difference pre-to-post-intervention on one item from the MDS-R. No other results from the MDS-R are reported.

3.5.4. Single-group design studies

Results were mixed for single-group design studies. Brandon et al. (2014) found a statistically significant decrease in the unadjusted frequency for the moral distress subscale '*Not in the patient's best interests*' using their modified two subscale version of the instrument. There were no significant differences when adjusted for discipline, setting, or years at institution. They did find a statistically significant difference in the adjusted analysis for intensity of work quality of life post-intervention. Monteverde (2016) hypothesized that those who scored lower levels of moral distress when considering ethically complex situations had greater moral resilience. Monteverde (2016) found a significant decrease in MDT responses for three of four purposely written vignettes following the educational intervention. Rushton et al. (2021) reported no change in moral distress pre-to-post the MEPRA intervention. Wocial et al. (2017) found that MDS-R scores 12-months after the implementation of multidisciplinary PEACE rounds were lower overall but not all scores were

statistically significant. Results showed physicians experienced a significantly improved score for one item *feel pressure to order what I consider to be unnecessary tests and treatments* and nurses experienced improved scores for three items: *initiate extensive life-saving actions when I think they only prolong death; work with nurses or other healthcare providers who are not as competent as the patient care requires*; and *witness diminished patient care quality due to poor team communication.*

3.5.5. Mixed methods research

Results in mixed methods studies were generally favorable with positively expressed themes in the qualitative analysis but with quantitative results that were contradictory or mixed. Chiafery et al. (2018) reported a statistically significant decrease in MDT scores post-intervention. Qualitative content analysis of text responses in the survey revealed generally positive themes: feeling supported after sharing feelings about moral distress with others, validation of experiences, ability to engage in perspective-taking and gaining new knowledge that would be useful in future morally distressing situations. Allen and Butler (2016) could not report inferential statistical analysis due to a small sample size. Three of the four nurses in the intervention phase reported unchanged MDS-R scores at 3 months post-intervention (Allen & Butler, 2016). Conversely, qualitative results demonstrated that participants initially felt as if they persevered through moral distress initially but utilized skills they had learned at 3-month follow up.

3.5.6. Quasi-experimental study

Beumer (2008) reported improvements in the percentage of participants who chose 'agree options' on a 5-point Likert scale (ranging from strongly agree to strongly disagree) following an educational workshop. However, the percentage improvements were not significant when compared to the control group that consisted of supplemental (float pool) staff. This suggests that the workshop had limited impact on the intervention group when compared to the control.

3.5.7. Randomized controlled trials

Results for randomized controlled trials were mixed, with longitudinal analysis showing improvement in moral distress outcome measures. For example, Saaedi et al. (2018) found no significant difference in mean scores between experimental and control groups who took the MDS-R immediately following their eight-week long intervention. Conversely, both Abassi et al. (2018) and Molazem et al. (2013) showed significant differences in mean moral distress outcome measures between groups 1 month after the intervention (Abassi et al., 2018) and 1 and 2- months after the intervention (Molazem et al., 2013).

4. Discussion

This review examined and evaluated interventions developed to mitigate moral distress. In this section, we will discuss the findings from the review. First, we will describe the methodological weaknesses identified in the included studies. Second, we discuss the difficulty of designing interventions to mitigate a complex phenomenon. Third, we highlight deficiencies in meeting expected reporting standards. Finally, we discuss the promising direction offered by educational interventions and suggest that interprofessional collaboration between ethicists and scientists might ensure that the identified limitations are overcome in future research.

A key finding from our review was that the vast majority of included articles lacked scientific rigor, and had significant threats to internal and external validity. Scientific rigor starts with choosing the appropriate design that will achieve a specified goal, such as the generation of new, generalizable knowledge from a research study or evaluating merit, worth, and importance in program evaluation. Randomized controlled trials (RCT) are the gold standard for testing the effectiveness of an intervention. There were only three RCTs in our analysis, two of which reported significant findings (Abassi et al., 2019; Molazem et al., 2013; Saaedi et al., 2018). Researchers were at risk for making a Type I or Type II error because they did not report an a priori sample size calculation nor did they complete a post hoc power analyses. Molazem et al. (2013) and Saeedi et al. (2018) did not report their allocation mechanism and sample sizes ranged from 30 per group (Abassi et al., 2018; Molazem et al., 2013) to 120 total participants (Saaedi et al., 2018). Saaedi et al. (2018) were the only authors to present a CONSORT diagram, intent-to-treat analysis, and single blinding (Eldridge et al., 2016).

The most common threat to internal validity was selection bias. The vast majority of studies used convenience samples of participants. For example, Robinson et al. (2014) may have selected highly motivated individuals who might be more likely to report high self-efficacy and knowledge post intervention. Participation in debriefing sessions was voluntary but self-selection could bias the findings because individuals who find it cathartic or helpful to discuss morally distressing experiences would attend and report a positive experience (Browning & Cruz, 2018; Fontenot & White, 2019). Whilst this is a threat to the validity of the research, debriefing may still have potential as an intervention given that some participants who chose to attend did report positive experiences (Browning & Cruz, 2018).

Perhaps the biggest problem we encountered in the studies we reviewed were found in those that used single-group designs. Of the studies we reviewed, eight authors used single-group designs. Wocial et al. (2017), Hamric & Epstein (2017), Robinson et al. (2014), and Vaclavik et al. (2018) used designs in which a comparator group was not appropriate. Of the single group design studies that had control groups, they were not always comparable because they provided either no intervention or minimally comparable interventions. For example, an 8-week narrative writing program compared with a control group who received nothing (Saeedi et al., 2018) and 10-hours of education compared to a control group who received no education (Beumer, 2008). In addition, much of the statistical analysis did not include paired pre-post test scores which meant the results were all aggregated and did not show change or effect of the intervention on moral distress. This threat is apparent in Allen & Butler (2016) who conducted their post surveys 3-months following their educational intervention; Beumer (2008) who completed questionnaires 7–10 weeks after the education intervention; Molazem et al. (2013) surveyed 1-month post and Brandon et al. (2014) who conducted their post surveys 20 months after the implementation of their pediatric quality of life program. A number of factors could therefore have impacted participants' responses.

The utility and frequency of single-group research designs used to evaluate the effect of interventions has been called into question by methodological experts due to the high risk of threat to both internal and external validity (Taylor and Asmundon, 2008; Knapp, 2016; Spurlock, 2018). The lack of a comparator group makes generalization of results beyond that single group impossible. Knapp (2016) suggests that this might be because researchers are more concerned with whether things get better, rather than why. Given that developing and testing interventions to address moral distress is still fairly new and is complex, Knapp's suggestion may be accurate. The studies included in this review have shown that, minimally, researchers should adhere to specific methods and reporting standards. However, it is likely that future work will require innovative methods and designs in order to test ethics interventions. We suggest that a promising direction would be to ex-

plore collaborations with empirical bioethicists or implementation scientists to apply methodologically rigorous methods to this field.

Undoubtedly, the complex and subjective nature of moral distress - in terms of both causes and effects - makes the design of interventions difficult. Moral distress can be caused by numerous moral constraints as captured on the MDS and MDS-R (Corley et al., 2001; Whitehead et al., 2014), and more recently authors have suggested other morally relevant causes of moral distress (Fourie, 2015; Campbell et al., 2016; Morley et al., 2020). Moral distress also manifests differently for individuals with different emotions ranging from frustration, anger and sorrow to physiological symptoms such as nausea, sleeplessness and migraines (Hanna, 2004; Wiegand & Funk, 2012). Due to this complexity, it is unlikely that one single intervention could prove to be effective to address all the possible causes and effects of moral distress. Presumably, it is the diffuse nature of moral distress that led some researchers to develop and test 'bundles' of interventions. Brandon et al. (2014), Vaclavik et al. (2018) and Rushton et al. (2021) implemented a number of different interventions all at one time. Brandon et al. (2014) provided education on end-of-life, debriefs, care planning and coordination services; Vaclavik et al. (2018) offered mindfulness techniques, debriefs, and a variety of supportive offerings; Rushton et al. (2021) combined education and mindfulness. Although bundled interventions may counter various causes and effects, the implementation of different interventions within one study increases threats to internal validity and provides limited data regarding the impact of each intervention within the bundle. It is likely that although interventions may need to be multifaceted in their approach, there still needs to be precision and rigor with implementation and measurement of effect. One suggestion would be to test one individual intervention to establish its efficacy before bundling interventions.

Other subjective phenomena such as pain, depression and anxiety can be somewhat successfully measured using a numeric rating scale - in addition to questionnaires - which perhaps highlights why the MDT seemed to be a more appropriate measure to test the efficacy of interventions. The MDT was designed to capture 'acute' moral distress whereas the MDS or MDS-R captures more global or 'chronic' moral distress. Two of the authors' optimized accuracy of their MD measure by using the MDT immediately post intervention (Chiafery et al., 2018; Fontenot & White, 2019).

An additional challenge of studying a complex concept that does not have one standard accepted definition is consistency between the underlying concept and the items or survey used to measure it. We found there was consistency regarding how authors defined moral distress when they employed previously validated instruments. Authors who modified or created their own measures did not state how they conceptualized or defined moral distress calling into question whether the instrument measures the underlying concept (Beumer, 2008; Brandon et al., 2014). Researchers who develop future instruments should be attentive to consistency between definition that underpins the measure and the measure being developed.

In addition to assessing research design and methods, we also noted a failure of authors to clearly elucidate their study design and report findings. Scientific rigor requires that researchers follow accepted reporting standards, such as those found in Equator Network (Equator Network, 2021). Unfortunately, many of the authors in the review did not follow this. Beumer (2008) do not state their study design so we based our classification of the design as quasi-experimental based on their description. Allen and Butler (2016) stated use of mixed methods but they did not meet a sufficient quantitative statistical analysis sample and failed to report a qualitative research design or analysis method. Robinson et al. (2014) initially suggested that they used a quasi-experimental research design but later stated that they were following a program evaluation format. Given that they described the program evaluation process in detail, we classified their design as program evaluation. Only Meziane et al. (2018) reported feasibility data in their pilot study, whereas other authors did not (Browning & Cruz, 2018; Fontenot & White, 2019; Wocial et al., 2017). Three author groups reported inferential statistical analysis of their moral distress outcome variable (Browning & Cruz, 2018; Fontenot & White, 2019; Meziane et al., 2018), but pilot studies are not powered to detect significant differences between groups (Kraemer et al., 2006; Thabane et al., 2010).

We also noted a level of variability with regard to the level of detail provided about each intervention in terms of the materials used, training/qualifications of facilitators, activities and modes of delivery. All of these elements are listed as information that should be included when describing an intervention according to TIDieR (Hoffman et al., 2014). Whilst this may in part be due to publishing restrictions, we would suggest that intervention studies should at minimum provide this information in a table that could be provided as supplementary material if required. This lack of detail in all studies meant that replicability was uniformly low. Of the RCTs, adherence to CONSORT reporting standards (Eldridge et al., 2016) was mixed yet eligibility criteria were clear in all three studies.

Although there appeared to be a wide-range of published interventions, clear themes did emerge regarding the types of interventions that were thought to mitigate moral distress. The groups of interventions reflected those suggested by (Burston and Tuckett, 2013) as either individualistic approaches aimed at improving understanding or coping mechanisms such as self-reflection or education, or collaborative approaches aimed at fostering an inter-professional environment and fostering group dialogue. We found that ethics education was the most prevalent and seemed to have some of the most promising results with four of the seven studies reporting statistically significant reductions in moral distress (Molazem et al., 2013; Abbasi et al., 2019; Monteverde, 2016; Robinson et al., 2014). These educational programs included communication techniques, role play, structured reflection and formulation of action plans.

Educational interventions varied in terms of theoretical foundations, content and duration which ranged from 30 minutes to 98 hours. Robinson et al. (2014) provided a program with the greatest duration (98 hours over 10-months) and their content was drawn from moral theory, the American Nurses Association Code of Ethics with interpretive Statements (2015) and clinical ethics competencies proposed by the American Society for Bioethics and Humanities (2009). Other authors used the *4As to Rise above Moral Distress* or Nathaniel's *Theory of Moral Reckoning.* Nathaniel (2006) used a grounded theory approach to elucidate the experiences and consequences of nurse moral distress but did not suggest this model could be used for building educational curricula or interventions and therefore its suitability is questionable. The *4 As* is potentially more suitable as an educational framework given that it provides steps for taking action when one experiences moral distress. For CERN, Robinson et al. (2014) drew from sources that are recognized for professional ethical standards which is much more theoretically sound. They also provided education and mentorship in ethical analysis and reasoning that would likely address the root cause of moral distress and equip participants with greater skills to address moral problems. This is supported by findings from Rathert et al. (2016) who surveyed nurses *(n=260)* at one US hospital and found that moral efficacy and voice – a nurse's ability to raise ethical concerns and to impact the outcome – mediated moral distress. Providing nurses with robust ethics education could build ethical confidence and skills, enabling them to engage in moral discussions and deliberation thus mitigating the negative effects of moral distress. Researchers would benefit from collaborating with education specialists to implement robust educational

interventions. Given the number of potential data points in educational interventions these might be better suited to program evaluation methods for assessment.

Intervention fidelity was an issue for all included studies because authors did not report on how they validated and tested their intervention. This was perhaps most notable in the facilitated discussions, which by nature of the intervention cannot be exactly replicated. Skillful facilitation is key to the success of facilitated discussion, and to an extent, also the provision of education. However, few authors reported on how facilitation occurred, by whom and the facilitation skills required. The duration of facilitated discussions varied from 30- 60 minutes with some sessions facilitated by a social worker and another a clinical ethicist. Although social workers bring highly valuable clinical skills to support the emotional and psychological experiences of moral distress, we would argue that relying on these skills alone ignores the moral dimensions of moral distress. As with educational interventions, an ethicist or an individual with advanced ethics knowledge is also required to sufficiently explore the ethical issues that are the root cause of moral distress.

The PEACE Rounds, Moral Distress Consultation Service (MDCS) and Nursing Ethics Huddles were all unit based interventions facilitated by clinical ethicists. These localized approaches seemed to enable facilitators to address moral distress caused by team dynamics, were sensitive to unit environments in which nurses usually work and were facilitated by individuals with expertise in managing ethical issues in clinical care. PEACE Rounds and Nursing Ethics Huddles were both found to reduce moral distress. The MDCS also enabled organizational challenges to be addressed by elevating unit concerns to leaders and influencing policy. Due to a lack of reporting regarding data collection and analysis, it is difficult to fully evaluate the impact of the MDCS on moral distress but this approach does highlight the possible utility of system-wide approaches. Ethicists and philosophers should consider working with researchers to establish and implement suitable research methods.

5. Limitations

Although we endeavored to include all relevant literature retrieved, some relevant studies may have been inadvertently missed. In the review retrieval process, only one author conducted relevance assessment. However, given that the first author has expertise in moral distress and in-depth knowledge of the pre-existing literature this was deemed reasonable from a practical standpoint. The first author had a low tolerance for borderline papers which were then reviewed by the research team for relevance. This is reflected in Fig. 1 as forty papers were reviewed by the team. Additional limitations of this review are the lack of quantitative comparison between studies because of the infeasibility and the moderate quality of some of the included papers. We decided to include quality improvement and program evaluation reports because there were a limited number of interventions reported and because we hope that researchers will use this review to improve future interventions and study designs.

6. Conclusions

Many of the studies included in this review that purport to describe interventions intended to address or reduce moral distress lacked scientific rigor. Methodologically, many of the studies were lacking because of single group designs, which are frequently used for pilot studies and lack generalizability of results (Knapp, 2016; Spurlock, 2018). When controls were utilized they were not adequately comparable. Many of the authors failed to adequately describe their intervention, likely because they did not adhere to reporting standards for interventions, and therefore their replicability is questionable. Effect sizes also could not be measured due to inadequate statistical power and analysis. Many of the researchers, for example those who offered yoga and mindfulness, failed to consider how they would address the underlying moral event that causes moral distress. These interventions arguably have benefit for mitigating psychological distress or stress but the authors do not report this as either an aim or a limitation, which raises concern about their understanding of moral distress. Future researchers should aim to focus clearly on the aspects of moral distress they are hoping to address and utilize precise measures rather than employing multipronged approaches that lack precision. Moral distress continues to be a difficult ethical phenomenon to address and conducting research using traditional scientific methods such as randomized controlled trials and interventional studies may arguably be unsuitable, or require thoughtful modification. Such modification would need to be rigorous, coherent and transparently justified. Researchers would likely benefit from collaborating with clinical ethicists who can assist in providing the required expertise to explore ethical issues in clinical practice, and with empirical bioethicists who can provide recommendations regarding ways to explore ethical issues empirically. We conclude by calling for more collaborative approaches to this field of research and the use of innovative methods that are rigorous and well-justified in their design in order to measure the effects of interventions to address moral distress.

Conflict of Interest

The authors have no conflicts of interest to declare

Funding

None associated with this work.

References

Abbasi, S., Ghafari, S., Shahriari, Shahgholian, N., 2019. Effect of moral empowerment program on moral distress in intensive care units. Nurs. Ethics 26 (5), 1494–1504. doi:10.1177/0969733018766576.

Allen, R., Judkins-Cohn, T., deVelasco, R., Forges, E., Lee, R., Clark, L., Procunier, M., 2013. Moral distress among healthcare professionals at a health system. JONA'S Healthcare Law. Ethics Regul. 15 (3), 111–118.

Allen, R., Butler, E., 2016. Addressing moral distress in critical care nurses: a pilot study. Int. J. Crit. Care Emerg. Med. 2 (1), 1–6.

American Association of Critical-Care Nurses, 2004. The 4 A's to Rise Above Moral Distress. AACN.

American Society for Bioethics and Humanities, 2009. Improving Competencies in Clinical Ethics Consultation: An Education Guide. ASBH. Glenview. IL..

American Nurses Association, 2015. Code of Ethics for Nurses with Interpretive Statements. MORAL DISTRESS: Sliverspring.

Beumer, C., 2008. The Effect of a Workshop on Reducing the Experience of Moral Distress in an Intensive Care Unit Setting. Educ. Dimension 27 (6), 263–267.

Brandon, D., Ryan, D., Sloane, R., Docherty, S.L., 2014. Impact of a pediatric quality of life program on providers' moral distress. MCN Am. J. Matern. Child Nurs. 39 (3), 189–197. doi:10.1097/NMC.0000000000000025.

Browning, E.D., Cruz, J.S., 2018. Reflective debriefing: a social work interventions addressing moral distress among ICU nurses. J. Social Work End-of-Life Palliative Care 14 (1), 44–72.

Burston, A.S., Tuckett, A.G., 2013. Moral distress in nursing: contributing factors, outcomes and interventions. Nurs. Ethics 20 (3), 312–324. doi:10.1177/0969733012462049.

Campbell, S.M., Ulrich, C.M., Grady, C., 2016. A broader understanding of moral distress. Am. J. Bioeth. 16 (12), 2–9.

Chiafery, M.C., Hopkins, P., Norton, S.N., Hodges Shaw, M., 2018. Nursing ethics huddles. J. Clin. Ethics 29 (3), 217–226.

Corley, MC., 1995. Moral Distress of Critical Care Nurses. Am. J. Crit. Care. 4 (4), 280–285.

Corley, MC., Elswick, RK., Gorman, M., Clor, T., 2001. Development and evaluation of a moral distress scale. J. Adv. Nurs. 33 (2), 250–256. doi:10.1046/j.1365-2648.2001.01658.x.

Eldridge, S.M., Chan, C.L., Campbell, M.J., Bond, C.M., Hopewell, S., Thabane, L., Lancaster, G.A.PAFS consensus group, 2016. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 355, i5239. doi:10.1136/bmj.i5239.

Equator Network, 2021 The UK EQUATOR Centre. https://www.equator-network.org/ Accessed on 19/03/2021.

Fontenot, N., White, K.A., 2019. Using evidence-based debriefing to combat moral distress in critical care nurses: a pilot project. J. Nursing Educ. Pract. 9 (1), 1–12.
Fourie, C., 2015. Moral distress and moral conflict in clinical ethics. Bioethics 29 (2), 91–97.
Hamric, A., 2012. Empirical research on moral distress: issues, challenges, and opportunities. HEC Forum 24, 39–49.
Hamric, A., Borchers, C., Epstein, E., 2012. Development and Testing of an Instrument to Measure Moral Distress in Healthcare Professionals. AJOB Prim. Res. 3 (2), 1–9.
Hamric, A., Epstein, E., 2017. A health system-wide moral distress consultation service: development and evaluation. HEC Forum 29 (2), 127–143.
Hanna, D., 2004. Moral distress: the state of the science. Res. Theory Nursing Pract. 18 (1), 73–93. doi:10.1891/rtnp.18.1.73.28054, PMID: 15083663.
Helft, P.R., Bledsoe, P.D., Hancock, M., Wocial, L.D., 2009. Facilitated ethics conversations a novel program for managing moral distress in bedside nursing staff. JONA'S healthcare law. Ethics Regul. 11 (1), 27–33.
Hoffmann, T.C., Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S.E., Dixon-Woods, M., McCulloch, P., Wyatt, J.C., Chan, A.W., Michie, S., 2014. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 348, g1687. doi:10.1136/bmj.g1687.
Howe, E., 2017. Fourteen Important Concepts Regarding Moral Distress. J. Clin. Ethics 28, 3–14.
Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Tianjing, L., Page, M.J., Welch, V.A, 2019. Cochrane Handbook for Systematic Reviews of Interventions, 2nd edition ed. Wiley-Blackwell, Chichester, UK.
Jameton, A., 1984. Nursing Practice: The Ethical Issues. Prentice Hall, Englewood Cliffs, NJ.
Johns, C., 2006. Engaging reflection in practice: a narrative approach. Blackwell Publishing, Oxford.
Knapp, T., 2016. Why Is the One-Group Pretest-Posttest Design Still Used? Clin. Nurs. Res. 25 (5), 467–472. doi:10.1177/1054773816666280.
Kraemer, H.C., Mintz, J., Noda, A., Tinklenberg, J., Yesavage, J.A., 2006. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch. Gen. Psychiatry 63 (5), 484–489. doi:10.1001/archpsyc.63.5.484.
Maffoni, M., Argentero, P., Giorgi, I., Hynes, J., Giardini, A., 2019. Healthcare professionals' moral distress in adult palliative care: a systematic review. BMJ Supp. Palliat. Care 9 (3), 245–254. doi:10.1136/bmjspcare-2018-001674.
Maiden, J., Georges, J.M., Connelly, C.D., 2011. Moral distress, compassion fatigue, and perceptions about medication errors in certified critical care nurses. Dimens. Crit. Care Nurs. 30, 339–345.
McAndrew, N.S., Leske, J., Schroeter, K., 2018. Moral distress in critical care nursing: the state of the science. Nurs. Ethics 25 (5), 552–570. doi:10.1177/0969733016664975.
Meziane, D., Ramirez-Garcia, M.P., Fortin, M.L., 2018. A reflective practice intervention to act on the moral distress of nurses providing end-of-life care on acute care units. Int. J. Palliat. Nurs. 24 (9), 444–451. doi:10.12968/ijpn.2018.24.9.444.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G.PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 21 (339), b2535. doi:10.1136/bmj.b2535.
Molazem, Z., Tavakol, N., Sharif, F., Keshavarzi, S., Ghadakpour, S., 2013. Effect of education based on the "4A Model" on the Iranian nurses' moral distress in CCU wards. J. Med. Ethics Hist. Med. 6 (5), 1–8.
Morley, G., 2018. What is moral distress in nursing and how should we repsond to it? Unpublished thesis. https://research-information.bris.ac.uk/ws/portalfiles/portal/175857763/Final_Copy_2018_09_25_Morley_G_PhD_Redacted.pdf
Morley, G., Bradbury-Jones, C., Ives, J., 2020. What is 'moral distress' in nursing? A feminist empirical bioethics study. Nurs. Ethics 27 (5), 1297–1314. doi:10.1177/0969733019874492, 2020..
Morley, G., Bradbury-Jones, C., Ives, J., 2021. Reasons to redefine moral distress: a feminist empirical bioethics analysis. Bioethics. 35 (1), 61–71. doi:10.1111/bioe.12783.
Monteverde, S., 2016. Caring for tomorrow's workforce: Moral resilience and healthcare ethics education. Nurs. Ethics 23 (1), 104–116.
Nathaniel, A., 2006. Moral Reckoning in Nursing. West. J. Nurs. Res. 28 (4), 419–438.
Neumann, J.L., Mau, L.W., Virani, S., Denzen, E.M., Boyle, D.A., Boyle, N.J., Dabney, J., De KeselLofthus, A., Kalbacker, M., Khan, T., Majhail, N.S., Murphy, E.A., Paplham, P., Parran, L., Perales, M.A., Rockwood, T.H., Schmit-Pokorny, K., Shanafelt, T.D., Stenstrup, E., Wood, W.A., Burns, L.J., 2018. Burnout, Moral Distress, Work-Life Balance, and Career Satisfaction among Hematopoietic Cell Transplantation Professionals. Biol. Blood Marrow Transplant. 24 (4), 849–860. doi:10.1016/j.bbmt.2017.11.015.
Nyholm, S., 2016. The normative and evaluative status of moral distress in health care contexts. Am. J. Bioeth. 16 (12), 17–19.
Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., Duffy, S. 2006. Guidance on the Condcut of Narrative Synthesis in Systematic Reviews: A Product from the ESRC Methods Programme. Lancaster University.
Rathert, C., May, D.R., Chung, H.S., 2016. Nurse moral distress: a survey identifying predictors and potential interventions. Int. J. Nurs. Stud. 53, 39–49. doi:10.1016/j.ijnurstu.2015.10.007.
Rhodes, R., Alfandre, D., 2007. A systematic approach to clinical moral reasoning. Clin. Ethics 2 (2), 66–70.
Robinson, E., Lee, S., Zollfrank, A., Jurchak, M., Frost, D., Grace, P., 2014. Enhancing moral agency: clinical ethics residency for nurses. Hastings Cent. Rep. 44 (5), 12–20. doi:10.1002/hast.353.
Rushton, C.H., Batcheller, J., Schroeder, K., Donohue, P., 2015. Burnout and resilience among nurses practicing in high-intensity settings. Am. J. Crit. Care. 24 (5), 412–420. doi:10.4037/ajcc2015291.
Rushton, C.H., Swoboda, S.M., Reller, N., Skarupski, K.A., Prizzi, M., Young, P.D., Hanson, G.C., 2021. Mindful ethical practice and resilience academy: equipping nurses to address ethical challenges. Am. J. Crit. Care. 30 (1), e1–e11. doi:10.4037/ajcc2021359.
Saeedi, S., Jouybari, L., Sanagoo, A., Vakili, M.A., 2018. The effectiveness of narrative writing on the moral distress of intensive care nurses. Nurs. Ethics 26 (7-8), 2195–2203.
Spurlock, D., 2018. The single-group, pre- and posttest design in nursing education research: it's time to move on. J. Nurs. Educ. 57 (2), 69–71. doi:10.3928/01484834-20180123-02.
Strech, D.S., Synofzik, M., Marckmann, G., 2008. Systematic reviews of empirical bioethics. J. Med. Ethics 34 (6), 472–477.
Taylor, S., Asmundon, G.J.G, 2008. Internal and external validity in clinical research. In: McKay, D. (Ed.), Handbook of research methods in abnormal and clinical psychology. Sage Publications, Los Angeles, pp. 23–34.
Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L.P., Robson, R., Thabane, M., Giangregorio, L, Goldsmith, C.H., 2010. A tutorial on pilot studies: the what, why and how. BMC Med. Res. Methodol. 10 (1), 1–10. doi:10.1186/1471-2288-10-1.
Tigard, D. 2018. Rethinking moral distress: conceptual demands for a troubling phenomenon affecting health care professionals. Med. Health Care Philos., 21 (4), 479–488. doi:10.1007/s11019-017-9819-5
Vaclavik, E., Staffileno, B.A., Carlson, E., 2018. Moral distress: using mindfulness-based stress reduction interventions to decrease nurse perceptions of distress. Clin. J. Oncol. Nurs. 22 (3), 326–332. doi:10.1188/18.
Wiegand, D., Funk, M., 2012. Consequences of clinical situations that cause critical care nurses to experience moral distress. Nurs. Ethics 19, 479–487.
Whitehead, P., Herbertson, R., Hamric, A., Epstein, E., Fisher, J., 2014. Moral Distress Among Healthcare Professionals: Report of an Institution-Wide Survey. J. Nurs. Scholarsh. 47 (2), 117–125. doi:10.1111/jnu.12115.
Wocial, L.D., Weaver, M.T., 2013. Development and psychometric testing of a new tool for detecting moral distress: the Moral Distress Thermometer. J. Adv. Nurs. 69 (1), 167–174. doi:10.1111/j.1365-2648.2012.06036.x.
Wocial, L.D., Ackerman, V., Leland, B., Benneywroth, B., Patel, V., Tong, Y., Nitu, M., 2017. Pediatric Ethics and Communication Excellence (PEACE) Rounds: Decreasing Moral Distress and Patient Length of Stay in the PICU. HEC Forum 29 (1), 75–91.