



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/131932/>

Version: Accepted Version

Article:

Johnson, J and Panagioti, M (2018) Interventions to Improve the Breaking of Bad or Difficult News by Physicians, Medical Students, and Interns/Residents: A Systematic Review and Meta-Analysis. *Academic Medicine*, 93 (9). pp. 1400-1412. ISSN: 1040-2446

<https://doi.org/10.1097/ACM.0000000000002308>

© 2018 by the Association of American Medical Colleges. This is an author produced version of a paper published in *Academic Medicine*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

ACADEMIC MEDICINE

Journal of the Association of American Medical Colleges

Uncomposed, edited manuscript published online ahead of print.

This published ahead-of-print manuscript is not the final version of this article, but it may be cited and shared publicly.

Author: Johnson Judith PhD, ClinPsyD; Panagioti Maria PhD

Title: Interventions to Improve the Breaking of Bad or Difficult News by Physicians, Medical Students, and Interns/Residents: A Systematic Review and Meta-Analysis

DOI: 10.1097/ACM.0000000000002308

ACCEPTED

Academic Medicine

DOI: 10.1097/ACM.0000000000002308

Interventions to Improve the Breaking of Bad or Difficult News by Physicians, Medical Students, and Interns/Residents: A Systematic Review and Meta-Analysis

Judith Johnson, PhD, ClinPsyD, and Maria Panagioti, PhD

J. Johnson is joint lecturer, School of Psychology, University of Leeds, Leeds, and Bradford Institute for Health Research, Bradford, United Kingdom; ORCID: <https://orcid.org/0000-0003-0431-013X>.

M. Panagioti is senior research fellow, National Institute for Health Research School for Primary Care Research, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom; ORCID: <https://orcid.org/0000-0002-7153-5745>.

Correspondence should be addressed to Judith Johnson, School of Psychology, University of Leeds, Lifton Place, Leeds LS29JT, UK; telephone: +44 (0)1133435719; e-mail: j.johnson@leeds.ac.uk; Twitter: @DrJTJohnson.

Supplemental digital content for this article is available at

<http://links.lww.com/ACADMED/A564>.

Funding/Support: This research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Yorkshire and Humber (www.clahrc-yh.nihr.ac.uk; reference: NIHR IS-CLA-0113-10020). In addition, the NIHR School for Primary Care Research and the NIHR Greater Manchester Patient Safety Translational Research Centre (Greater Manchester PSTRC) supported the time and facilities of M.P. The research team members were independent from the funding agency.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Disclaimer: The views expressed in this report are those of the authors, and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care. The funders had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; and the preparation, review, or approval of the manuscript.

ACCEPTED

Abstract

Purpose

To assess the effectiveness of news delivery interventions to improve observer-rated skills, physician confidence, and patient-reported depression/anxiety.

Method

MEDLINE, EMBASE, CINAHL, PsycINFO, and Cochrane Register of Controlled Trials databases were searched from inception to September 5, 2016 (updated February 2017). Eligible studies included randomized controlled trials (RCTs), non-RCTs, and controlled before-after studies of interventions to improve the communication of bad or difficult news by physicians, medical students, and residents/interns. The EPOC risk of bias tool was used to conduct a risk of bias assessment. Main and secondary meta-analyses examined the effectiveness of the identified interventions for improving observer-rated news delivery skills and for improving physician confidence in delivering news and patient-reported depression/anxiety, respectively.

Results

Seventeen studies were included in the systematic review and meta-analysis, including 19 independent comparisons on 1,322 participants and 9 independent comparisons on 985 participants for the main and secondary (physician confidence) analyses (mean [SD] age = 35 [7] years; 46% male), respectively. Interventions were associated with large, significant improvements in observer-rated news delivery skills (19 comparisons: standardized mean difference [SMD] = 0.74, 95% CI = 0.47–1.01) and moderate, significant improvements in physician confidence (9 comparisons: SMD = 0.60, 95% CI = 0.26–0.95). One study reported intervention effects on patient-reported depression/anxiety. The risk of bias findings did not influence the significance of the results.

Conclusions

Interventions are effective for improving news delivery and physician confidence. Further research is needed to test the impact of interventions on patient outcomes and determine optimal components and length.

ACCEPTED

Physicians frequently break bad or difficult news to patients. Research into the delivery of bad or difficult news originated in oncology services, where it was found that news communication practices can have a strong and lasting impact on patients' subsequent symptoms of depression, anxiety, and post-traumatic stress disorder and can influence their treatment choices.¹⁻⁵ Studies have since investigated the delivery of bad and difficult news in a range of health care settings, including pediatrics,⁶ emergency medicine,^{7,8} and obstetrics services.^{9,10} Together, this body of research has identified several challenges that physicians may face in scenarios where they have to deliver bad or difficult news, such as when the news occurs suddenly and without warning (e.g., in emergency settings), when there is limited time for physicians to prepare to deliver the news (e.g., in obstetric ultrasound settings), or when the news itself is uncertain because the diagnosis or prognosis is unclear. It has also highlighted the negative impact that these events can have on the physicians involved, including increased stress and burnout.^{8,11}

A range of interventions that aim to improve the communication skills and confidence of physicians in delivering bad or difficult news have been described. These interventions vary in length and format, but share some similar components; for example, most include elements of didactic teaching, role-playing or simulation,⁹ group discussions,¹² or the viewing of instructional videos.¹³ The interventions are often designed to enhance fidelity to existing guiding frameworks for bad or difficult news delivery. SPIKES¹⁴ is the most widely used of these frameworks; it proposes six steps, from which the acronym is derived, to improve news delivery events. These steps are (1) setting up the interview, (2) assessing the patient's perception of the situation, (3) obtaining the patient's invitation to deliver the news, (4) giving knowledge and information to the patient, (5) addressing the patient's emotions empathically, and (6) providing a summary and discussing prognosis and treatment options.¹⁴ Similarly, the more recently proposed SHARE¹⁵ protocol suggests that health care staff

should follow four steps, which taken together form the acronym. These steps are (1) create a supportive environment, (2) consider how to deliver the news, (3) discuss additional information that patients would like to know, and (4) provide reassurance and emotional support.¹⁵

Divergent methods have been used to evaluate the effectiveness of these interventions, but the most common practice has been the rating of a participant's news delivery skills in a simulated exercise by an observer.^{12,13,16} Other practices have included measuring physician confidence in breaking bad news¹⁷ and gathering information on the patient's experience.¹⁵ In the United Kingdom, National Institute for Health and Care Excellence (NICE) guidelines for miscarriages state that staff should be trained in delivering bad or difficult news,¹⁸ and the NICE list of quality statements state all National Health Service staff should have competency in communication skills.¹⁹ Similarly, the US Preventive Services Task Force recommendations for prostate and lung cancer screenings emphasize the importance of communicating with patients.^{20,21} However, news delivery interventions are not routinely implemented and patients report poor satisfaction with this aspect of health care.^{22,23} Thus, there is a need to understand whether formal training interventions to improve the delivery of bad or difficult news by health care professionals are effective.

To address this need, we conducted the first systematic review and meta-analysis of studies that have evaluated bad or difficult news delivery interventions. Our first objective was to assess the effectiveness of these interventions for improving news delivery skills in physicians, medical students, or residents/interns, as rated by an observer such as a researcher, instructor, or standardized patient (an individual who is trained to role play a patient in a standardized format). Our second objective was to assess the effectiveness of these interventions for improving physician confidence in news delivery. Our third objective was to assess the impact of the interventions on improving patient-reported depression and/or

anxiety. Our fourth objective was to compare observer ratings of news delivery skills for SPIKES-based interventions with observer ratings of news delivery skills for interventions based on any other or no framework.

Method

Our review followed the PRISMA statement²⁴ (see eTable 1 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>). We prospectively registered our protocol (see eMethods 1 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>) with PROSPERO (CRD42016045892).

Search strategy and data sources

We searched five electronic bibliographic databases—MEDLINE, EMBASE, CINAHL, PsycINFO, and Cochrane Register of Controlled Trials—from inception to September 5, 2016. We searched the same five databases in February 2017 to update our results. Our search strategy included combinations of two key blocks of terms (bad or difficult news and intervention) using a combination of medical subject headings (MeSH terms) and text words (see eMethods 2 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>). We scanned the reference lists of eligible studies to identify additional potentially includable studies.

Eligibility criteria

Studies that met the criteria in each of the areas listed below were eligible for inclusion in the systematic review and meta-analysis:

- **Population:** Studies that looked at physicians, medical students, or residents/interns working in any primary, secondary, or intensive clinical health care settings were included.
- **Intervention:** Studies with interventions that were designed to improve the delivery of bad or difficult news were included. These mainly included person-

directed training programs for improving news communication skills using real or simulated scenarios with or without an underlying theoretical framework (e.g., SPIKES).

- **Comparison:** Studies with any type of control group (e.g., no intervention, alternative intervention, wait list) were included.
- **Outcomes:** Studies in which the outcome of news delivery skill as rated by an observer (e.g., researcher, standardized patient) was measured were included. Where more than one news delivery skill metric was reported, we used the scale pertaining most closely to overall news delivery skill. Studies in which the outcomes of physician confidence in bad or difficult news delivery or patient-reported depression and/or anxiety were measured were also included.
- **Design:** Studies that had quantitative intervention designs, such as randomized controlled trials (RCTs), non-randomized controlled trials (nRCTs), controlled before-after studies (CBAs), and interrupted time series design studies (ITSs), as outlined in the Cochrane handbook, were included.²⁵
- **Context:** Studies conducted in any health care or educational setting that were English language and in peer-reviewed journals were included.

Exclusion criteria

Studies of communication interventions that did not assess outcomes relevant to breaking bad or difficult news were excluded. Studies that tested news delivery skills for “good,” “neutral,” or a range of news types, rather than focusing on “difficult,” “bad,” or “negative” news were also excluded, as were non-English language papers and gray literature.

Study selection

We exported the search results from each database and reference list scanning to Endnote version 7.1 (Clarivate Analytics, Philadelphia, Pennsylvania) and removed duplicates. We

undertook study selection in two stages. In the first stage, we screened the titles and abstracts of identified studies, and in the second stage, we accessed the full texts of the remaining studies to further screen them according to the eligibility criteria. We (J.J. and M.P.) independently reviewed a proportion of the titles and abstracts (10%) and our interrater reliability was high ($k = 1.0$). J.J. screened the remaining titles and abstracts. J.J. and M.P. completed full-text screenings for all eligible articles. We resolved any disagreements (which were rare) by discussion.

Data extraction

We devised a data extraction form in Excel 2010 (Microsoft Inc., Redmond, Washington) and piloted it on five randomly selected studies. We extracted quantitative data for the meta-analysis to a separate Excel file. J.J. undertook data extraction, with 10% of articles independently extracted by M.P. as well. We resolved any disagreements by discussion. We extracted the following descriptive information from eligible studies:

- Study: research design, recruitment method, and content of the control condition;
- Participants: sample size, age, gender, discipline, and setting;
- Intervention: content of the intervention, delivery format (group or individual), theoretical underpinning, and measurement time points; and
- Outcomes: observer-rated news delivery skill, physician confidence in bad or difficult news delivery, and patient-reported depression and/or anxiety.

Risk of bias assessment

We used the Effective Practice and Organisation of Care (EPOC) risk of bias tool²⁶ to conduct a critical appraisal, as it is appropriate for use across all different types of intervention designs, as described in the Cochrane handbook,²⁵ including RCTs, nRCTs, CBAs, and ITSs. The EPOC tool contains nine standardized criteria, which are each rated on a three-point scale (where 0 = low risk, 1 = unclear risk, and 2 = high risk). We considered

studies that received a low risk score across at least six of the nine criteria to be less susceptible to risk of bias.

Data analysis

We synthesized our results using meta-analysis. We used standardized mean difference (SMD) as the effect size to pool results across the studies. We calculated SMDs and associated confidence intervals (CIs) for the news delivery outcomes of all the studies, and we pooled the results using the metaan command in Stata 15 (StataCorp LLC, College Station, Texas).²⁷ The main meta-analysis examined the effectiveness of the identified interventions for improving news delivery skills as rated by an observer (e.g., researcher, standardized patient). Secondary meta-analyses examined the effectiveness of the interventions for improving physician confidence in delivering bad or difficult news and for improving patient-reported depression and/or anxiety. When studies collected data at more than one follow-up assessment point, we used the first assessment point following the intervention. Pre-specified subgroup analyses²⁸ tested the effectiveness of interventions (via observer ratings of news delivery skills) for whether they were based on the SPIKES framework versus any other or no framework. We performed sensitivity analyses to examine whether results were maintained when only those studies with low risk of bias scores were included.

We used a random effects model to account for heterogeneity in all analyses. We assessed heterogeneity with the I^2 statistic and associated CIs. Conventionally, I^2 values of 25%, 50%, and 75% indicate low, moderate, and high heterogeneity, respectively.²⁹ We inspected the symmetry of funnel plots (asymmetry indicates publication bias), and we conducted the Egger's test of small-study effects to quantify observations in the funnel plots.³⁰ We constructed funnel plots using the metafunnel command³¹ and performed Egger's test using the metabias command in Stata 15.³²

Results

Our search strategy yielded 3,206 records (see Figure 1). An additional 41 records were identified via reference list scanning. Once we removed duplicates, we screened the titles and abstracts of the remaining 2,270 studies; 71 of these were retained for full-text screening. (For a list of the excluded studies, see eMethods 3 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>.) Seventeen studies were included in the systematic review and meta-analysis,^{9,13,15,33-47} with one study being reported in two papers,^{41,42} which we include separately in Appendix 1 but otherwise treat as a single study in the analyses and reporting. When studies were eligible for the systematic review but the necessary outcome data for the meta-analysis were not reported, we contacted the study authors for this information. However, there were two papers that were eligible for the review but were not ultimately included because we were not able to gather the information needed for the meta-analysis.^{48,49}

Characteristics of the studies and participants

Study characteristics are presented in Appendix 1. These included 19 independent comparisons on 1,322 participants for the main analysis of observer-rated news delivery skill^{9,13,15,33-35,37-41,43-47} and 9 independent comparisons on 985 participants for the secondary analysis of physician-reported confidence in news delivery^{36,37,40,42,44,45} (mean [SD] age = 35 [7] years). There was only one study reporting intervention effects on patient-reported depression and/or anxiety.¹⁵ The majority of studies took place in the United States (n = 10). One study was conducted in each of the following: Hong Kong,³⁶ Israel,³⁴ Japan,¹⁵ Puerto Rico,¹³ the United Kingdom,⁴³ and Belgium (this study was reported on in two papers^{41,42}). One study was conducted in both Germany and The Netherlands.⁴⁶ Overall, the proportion of male to female participants was approximately equal (46% male).

Five studies were conducted with medical students, 1 study was conducted with medical interns,³⁸ and 11 studies were conducted with physicians or residents. Most studies used a convenience sampling method, and the main eligibility criteria were being a medical student, intern/resident or physician working or training in a specific setting (data not shown). Studies were conducted in a range of settings, including oncology (n = 3), pediatrics (n = 3), university hospitals (n = 3), palliative care (n = 2), obstetrics (n = 2), primary care (n = 1),³⁴ hospital outpatient departments (n = 1),⁴⁶ and intensive care (n = 1).⁴³ The settings of two studies were unclear. Studies measured the outcome of interventions using a variety of tools, but 16 measured news delivery skills using an observer-rated measure, while 6 measured participant-reported news delivery self-confidence and 1 measured patient-reported depression and/or anxiety (see eTable 2 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>).¹⁵

Characteristics of the interventions

A range of theoretical models were used in the studies, but the single most commonly used model was SPIKES, which was cited in seven studies (see Appendix 1). Interventions varied in length; whereas some were delivered as a stand-alone intervention (n = 9), others were delivered in the context of a broader medical or communications training course (n = 8). The specific length of interventions delivered as part of wider training courses was not stated in all studies. However, the majority described interventions that lasted less than one working day, with the shortest intervention being 10 minutes.⁹ All were participant-directed interventions, and the majority involved a simulation exercise or role-play with feedback. None of the studies tested the impact of making service-level changes (e.g., the introduction of patient information leaflets) on bad or difficult news delivery. Two studies had more than one intervention condition to investigate the impact of the number and type of intervention components.^{37,38}

Risk of bias characteristics

For risk of bias characteristics, see eFigure 1 (in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>). Seven studies were RCTs, 2 were nRCTs where participants were allowed to switch their allocation after randomization if they were unable to attend their allocated condition, and 8 were CBAs. Five studies were scored as low on six or seven of the nine risk of bias criteria, indicating low overall risk of bias; 8 studies were scored as low on four or five of the criteria, indicating moderate overall risk of bias; and 5 studies were scored as low on two or three of the criteria, indicating high overall risk of bias.

Main and secondary meta-analyses

Interventions were associated with large, significant improvements in observer-rated news delivery skills (19 comparisons: SMD = 0.74, 95% CI = 0.47 to 1.01; $I^2 = 84%$, 95% CI = 40% to 98%; Figure 2), and moderate, significant improvements in physician confidence in news delivery (9 comparisons: SMD = 0.60, 95% CI = 0.26 to 0.95; $I^2 = 88%$, 95% CI = 43% to 99%; Figure 3). However, the heterogeneity between the studies was high in both analyses (as indicated by the I^2 statistic [shown above]).

As only one study investigated patient-reported depression and/or anxiety,¹⁵ a meta-analysis of this outcome was not possible; however, we did calculate the individual effect size in this study. The intervention group was associated with small, significant reductions in patient-reported depression and/or anxiety symptoms, as compared with the control group (SMD = -0.17, 95% CI = -0.33 to -0.01).

Subgroup analyses for use of SPIKES framework

The pooled effect size of interventions on observer-rated news delivery skills was very large and significant across the studies that used SPIKES (7 comparisons: SMD = 1.14, 95% CI = 0.63 to 1.66; $I^2 = 92%$, 95% CI = 40% to 100%) and was moderate and significant across the studies that used any other or no framework (12 comparisons: SMD = 0.52, 95% CI = 0.21 to

0.83; $I^2 = 70\%$, 95% CI = 20% to 90%; see eFigure 2 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>). The effects of SPIKES-based interventions were significantly larger than those of other interventions ($Q = 4.23$, $P = .04$).

The pooled effect size of interventions on physician confidence was large and significant across the studies that used SPIKES (5 comparisons: SMD = 0.70, 95% CI = 0.14 to 1.26; $I^2 = 94\%$, 95% CI = 40% to 100%) and was medium and significant across the studies that used any other or no framework (4 comparisons: SMD = 0.47; 95% CI = 0.25 to 0.69; $I^2 = 0\%$, 95% CI = 0% to 1%; see eFigure 3 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>). However, this difference was not significant ($Q = 0.98$, $P = .69$).

Sensitivity analysis

The treatment effect derived by studies at lower overall risk of bias (i.e., those scoring low on six or more of the nine individual risk of bias criteria) showed a very large effect size for observer-rated news delivery skills (5 comparisons: SMD = 1.18, 95% CI = 0.40 to 1.96; $I^2 = 92\%$, 95% CI = 40% to 100%; see eFigure 4 in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>), suggesting the significance of results was not explained by risk of bias findings. Only one study scoring low on six or more of the risk of bias criteria included an item on physician confidence,⁴⁰ so no sensitivity analysis could be performed for this outcome.

Small-study bias

No evidence of funnel plot asymmetry, which might indicate publication bias, was identified for the main analysis focused on observer-rated news delivery skills (Egger's test: regression intercept = -0.24, in Supplemental Digital Appendix 1 at <http://links.lww.com/ACADMED/A564>).

Discussion

Summary of the main findings

This systematic review and meta-analysis found that interventions for improving the delivery of bad or difficult news were associated with significant, large improvements in observer-rated news delivery skills, and significant, moderate improvements in physician confidence in news delivery. The evidence base was too limited to test whether interventions improved patient-reported depression and/or anxiety. Interventions based on the SPIKES framework for bad or difficult news delivery were associated with very large and significant improvements in observer-rated news delivery skills, whereas interventions based on any other or no framework were associated with moderate and significant improvements.

Strengths and limitations

Strengths of this systematic review and meta-analysis included a pre-specified protocol registered on the PROSPERO database and inclusion of the EPOC tool to estimate risk of bias. We also endeavored to address the two most important threats of meta-analysis, which are publication bias and heterogeneity. We assessed publication bias using formal statistical tests, with no such bias indicators being observed in the performed tests.⁵⁰ A limitation of the meta-analysis was that we did find the heterogeneity between the studies to be high because the included studies differed considerably on a range of factors, such as the length of follow-up, outcome measures, and the content and length of the intervention. Most interventions lasted a day or less, with one being only 10 minutes.⁹ To account for this heterogeneity, we applied random effects models in all analyses and conducted pre-specified subgroup and sensitivity analyses (within the limits of power).^{51,52} The meta-analysis was also limited by its focus on outcomes immediately post-intervention, which prevented us from drawing conclusions about the long-term effects of interventions.

Comparison with previous meta-analyses

This is the first systematic review and meta-analysis of interventions to improve the delivery of bad or difficult news by physicians, medical students, and interns/residents. However, our findings are in line with two previous meta-analyses of broader communication skills training, which were restricted to oncology settings.^{53,54} A Cochrane review and meta-analysis of RCTs found that communication skills training interventions were effective for improving open-ended questioning and empathy in oncology clinicians.⁵³ Similarly, another meta-analysis of communication skills training in oncology found a moderate effect between communication skills training and positive communication behaviors.⁵⁴ This meta-analysis also reported that more extensive training resulted in greater improvements than shorter interventions and suggested a trend toward interventions leading to improved patient outcomes.⁵⁴ As with these previous meta-analyses, our meta-analysis found that a wide range of interventions have been used to train health care practitioners in communication skills, and there is no consensus regarding best practices. Our meta-analysis expands on these previous studies by focusing on one aspect of communication, bad or difficult news delivery, which is known to particularly influence the emotional well-being of both patients and physicians,^{1,2,8,55} and by finding evidence of effectiveness in a wider range of health care settings. However, our meta-analysis found comparatively larger overall effect sizes for the main and secondary outcomes (observer-rated news delivery skills and physician self-confidence). This may have been achieved by the tighter focus on news delivery interventions in comparison with generic communication interventions, as news delivery interventions are content-specific and which situations they should be applied to is clearer. This clarity may increase the likelihood of accurate and effective use by physicians over the long term and, therefore, support the use of news delivery interventions over more generic communication interventions.

Implications for clinicians, policymakers, and researchers

The importance of physicians being proficient in communication skills, particularly bad or difficult news delivery skills, is highlighted in guidance from both NICE (United Kingdom) and the US Preventive Task Force.^{18–21} However, research from a range of health care settings suggests a majority of patients are dissatisfied with this element of care.^{22,23} Our results suggest that news delivery interventions are effective in improving news delivery skills and indicate that implementation of such interventions could improve adherence to guidelines for communication in health care. Our results also suggest that interventions are more effective when based on the SPIKES framework for news delivery.¹⁴ This framework was initially developed in oncology but studies included in the meta-analysis applied it to pediatrics,^{13,35,44} obstetrics,^{9,40} and primary care,³⁴ suggesting that it can be adapted for a range of settings. It should be noted, though, that only limited information was available regarding the extent to which curriculum developers within each individual study relied on specific theoretical models, thus, this finding should be viewed tentatively. Our results also suggest that such interventions may be beneficial for the well-being of physicians. Breaking bad or difficult news is often experienced as stressful by physicians,^{8,42,55} but our meta-analysis suggests that interventions can significantly enhance physician confidence in this aspect of health care, which may reduce stress in relation to news delivery events and support physician well-being.

Our meta-analysis highlights two main limitations in the current evidence base, which could be addressed in future research. First, only one of the included studies measured patient outcomes (in this case, patient-reported depression and/or anxiety),¹⁵ and there is a need for trials to test whether interventions in news delivery skills translate to improved patient satisfaction and reduced patient depression and/or anxiety following news delivery events.

The one study that did test this found small, significant reductions in depression and/or

anxiety for patients seeing physicians trained in news delivery;¹⁵ while this is promising, it also indicates that such studies may require large sample sizes to detect effects. Second, there is a need to further test individual components of news delivery interventions to identify which of these actively impact skills improvement. Only two studies included in the meta-analysis had more than one intervention condition,^{37,38} which prevented any sub-group analysis exploring whether adding more learning elements increased outcomes. Further research could also examine the optimal length of an intervention for producing positive change, as parsimonious interventions are increasingly desired in resource-limited health systems.

Summary

This systematic review and meta-analysis suggests that interventions for improving the breaking of bad or difficult news are effective for improving news delivery skills in physicians, medical students, and interns/residents and physician confidence in news delivery. Improvements were greater when interventions were based on the SPIKES framework. Further research is needed to test the impact of interventions on patient outcomes and to determine the optimal components and length.

References

1. Fallowfield LJ, Baum M, Maguire G. Effects of breast conservation on psychological morbidity associated with diagnosis and treatment of early breast cancer. *BMJ*. 1986;293:1331-1334.
2. Schofield PE, Butow PN, Thompson JF, Tattersall M, Beeney L, Dunn S. Psychological responses of patients receiving a diagnosis of cancer. *Annals of Oncology*. 2003;14:48-56.
3. Pruyn J, Rijckman R, Van Brunschot C, Van den Borne H. Cancer patients' personality characteristics, physician-patient communication and adoption of the Moerman diet. *Social Science & Medicine*. 1985;20:841-847.
4. Roberts CS, Cox CE, Reintgen DS, Baile WF, Gibertini M. Influence of physician communication on newly diagnosed breast patients' psychologic adjustment and decision-making. *Cancer*. 1994;74(1 Suppl):336-341.
5. Mager WM, Andrykowski MA. Communication in the cancer 'bad news' consultation: Patient perceptions and psychological adjustment. *Psycho-Oncology*. 2002;11:35-46.
6. Farrell M, Ryan S, Langrick B. 'Breaking bad news' within a paediatric setting: An evaluation report of a collaborative education workshop to support health professionals. *Journal of Advanced Nursing*. 2001;36:765-775.
7. Min AA, Spear-Ellinwood K, Berman M, Nisson P, Rhodes SM. Social worker assessment of bad news delivery by emergency medicine residents: A novel direct-observation milestone assessment. *Internal and Emergency Medicine*. 2016;11: 843-852.
8. Shaw JM, Brown RF, Dunn SM. A qualitative study of stress and coping responses in doctors breaking bad news. *Patient Education and Counseling*. 2013;91:243-248.

9. Karkowsky CE, Landsberger EJ, Bernstein PS, et al. Breaking bad news in obstetrics: A randomized trial of simulation followed by debriefing or lecture. *The Journal of Maternal-Fetal & Neonatal Medicine*. 2016;29:3717-3723.
10. Johnson J, Johnson O, Heyhoe J, Fielder C, Dunning A. Parent experiences and preferences when dyssemia is identified during the prenatal and perinatal periods: A qualitative study into family nursing care for rare diseases. *J Family Nurs*. 2018. In press.
11. Dosanjh S, Barnes J, Bhandari M. Barriers to breaking bad news among medical and surgical residents. *Medical education*. 2001;35:197-205.
12. Bradley CT, Webb TP, Schmitz CC, Chipman JG, Brasel KJ. Structured teaching versus experiential learning of palliative care for surgical residents. *The American Journal of Surgery*. 2010;200:542-547.
13. Silva D. A competency-based communication skills workshop series for pediatric residents. *Boletín de la Asociación Médica de Puerto Rico*. 2008;100:8-12.
14. Baile WF, Buckman R, Lenzi R, Glober G, Beale EA, Kudelka AP. SPIKES—A six-step protocol for delivering bad news: application to the patient with cancer. *The oncologist*. 2000;5:302-311.
15. Fujimori M, Shirai Y, Asai M, Kubota K, Katsumata N, Uchitomi Y. Effect of communication skills training program for oncologists based on patient preferences for communication when receiving bad news: A randomized controlled trial. *Journal of Clinical Oncology*. 2014;32:2166-2172.
16. Colletti L, Gruppen L, Barclay M, Stern D. Teaching students to break bad news. *The American Journal of Surgery*. 2001;182:20-23.

17. Shaw DJ, Davidson JE, Smilde RI, Sondoozi T, Agan D. Multidisciplinary team training to enhance family communication in the ICU. *Critical Care Medicine*. 2014;42:265-271.
18. National Institute for Health and Care Excellence: Diagnosis and Initial Management in Early Pregnancy of Ectopic Pregnancy and Miscarriage (CG154). London, UK: National Institute for Health and Care Excellence; 2012.
19. National Institute for Health and Clinical Excellence. Patient Experience in Adult NHS Services: Improving the Experience of Care for People Using Adult NHS Services (CG138). London, UK: National Institute for Health and Care Excellence; 2012.
20. US Preventive Services Task Force. Prostate Cancer: Screening. <https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/prostate-cancer-screening>. Published May 2012. Accessed April 11, 2018.
21. US Preventive Services Task Force. Lung Cancer: Screening. <https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/lung-cancer-screening>. Published December 2013. Accessed April 11, 2018.
22. Johnson J, Adams-Spink G, Arndt T, Wijeratne D, Heyhoe J, Taylor P. Providing family-centered care for rare diseases in maternity services: Parent satisfaction and preferences when screening identifies dysmelia. *Women and Birth*. 2016;29:e99–e104.
23. Seifart C, Hofmann M, Bär T, Riera Knorrenschild J, Seifart U, Rief W. Breaking bad news—What patients want and what they get: Evaluating the SPIKES protocol in Germany. *Annals of Oncology*. 2014;25:707-711.
24. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*. 2009;151:264-269.

25. Deeks J, Higgins J, Altman D. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. London, UK: The Cochrane Collaboration; 2011.
26. Effective Practice and Organisation of Care. Suggested risk of bias criteria for EPOC reviews. Oslo, Norway: Norwegian Knowledge Centre for the Health Services; 2014.
27. Kontopantelis E, Reeves D. metaan: Random-effects meta-analysis. *Stata Journal*. 2010;10:395-407.
28. Deeks J, Higgins J, Altman D. Undertaking subgroup analyses. In: Deeks J, Higgins J, Altman D. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. London, UK: The Cochrane Collaboration; 2011.
29. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557-560.
30. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-634.
31. Sterne JA, Harbord RM. Funnel plots in meta-analysis. *Stata Journal*. 2004;4:127-141.
32. Harbord RM, Harris RJ, Sterne JA. Updated tests for small-study effects in meta-analyses. *Stata Journal*. 2009;9:197.
33. Alexander SC, Keitz SA, Sloane R, Tulskey JA. A controlled trial of a short course to improve residents' communication with patients at the end of life. *Academic Medicine*. 2006;81:1008-1012.
34. Amiel GE, Ungar L, Alperin M, Baharier Z, Cohen R, Reis S. Ability of primary care physician's to break bad news: A performance based assessment of an educational intervention. *Patient Education and Counseling*. 2006;60:10-15.
35. Attar MA, Hernandez ME, Mullan PB, Tang TS, Haftel H. Pediatric residents' competency in communicating bad news and eliciting spiritual needs: Development

- and evaluation of a skills-based curriculum. *Journal of Neonatal-Perinatal Medicine*. 2010;3:177-185.
36. Betson C, Fielding R, Wong G, Chung S, Nestel D. Evaluation of two videotape instruction programmes on how to break bad news—For Cantonese-speaking medical students in Hong Kong. *Journal of Audiovisual Media in Medicine*. 1997;20:172-177.
 37. Bowyer MW, Hanson JL, Pimentel EA, et al. Teaching breaking bad news using mixed reality simulation. *Journal of Surgical Research*. 2010;159:462-467.
 38. Daetwyler CJ, Cohen DG, Gracely E, Novack DH. eLearning to enhance physician patient communication: A pilot test of “doc. com” and “WebEncounter” in teaching bad news delivery. *Medical Teacher*. 2010;32(9):e381-e390.
 39. Gorniewicz J, Floyd M, Krishnan K, Bishop TW, Tudiver F, Lang F. Breaking bad news to patients with cancer: A randomized control trial of a brief communication skills training module incorporating the stories and preferences of actual patients. *Patient Education and Counseling*. 2017;100:655-666.
 40. Marko EK, Buery-Joyner SD, Sheridan MJ, Nieves K, Khoury AN, Dalrymple JL. Structured teaching of early pregnancy loss counseling. *Obstetrics & Gynecology*. 2015;126(Suppl 4):1S-6S.
 41. Merckaert I, Liénard A, Libert Y, et al. Is it possible to improve the breaking bad news skills of residents when a relative is present? A randomised study. *British Journal of Cancer*. 2013;109:2507-2514.
 42. Meunier J, Merckaert I, Libert Y, et al. The effect of communication skills training on residents’ physiological arousal in a breaking bad news simulated task. *Patient Education and Counseling*. 2013;93:40-47.

43. Morton J, Blok G, Reid C, van Dalen J, Morley M. The European Donor Hospital Education Programme (EDHEP): Enhancing communication skills with bereaved relatives. *Anaesthesia and Intensive Care*. 2000;28:184-190.
44. Nellis ME, Howell JD, Ching K, Bylund C. The use of simulation to improve resident communication and personal experience at end-of-life care. *Journal of Pediatric Intensive Care*. 2017;6:91-97.
45. Szmuiłowicz E, El-Jawahri A, Chiappetta L, Kamdar M, Block S. Improving residents' end-of-life communication skills with a short retreat: A randomized controlled trial. *Journal of Palliative Medicine*. 2010;13:439-452.
46. Wijnen-Meijer M, ten Cate O, van der Schaaf M, Burgers C, Borleffs J, Harendza S. Vertically integrated medical education and the readiness for practice of graduates. *BMC Medical Education*. 2015;15:229.
47. Vetto JT, Elder NC, Toffler WL, Fields SA. Teaching medical students to give bad news: Does formal instruction help? *Journal of Cancer Education*. 1999;14:13-17.
48. Pekmezaris R, Walia R, Nouryan C, et al. The impact of an end-of-life communication skills intervention on physicians-in-training. *Gerontology & Geriatrics Education*. 2011;32:152-163.
49. Bowyer MW, Rawn L, Hanson J, et al. Combining high-fidelity human patient simulators with a standardized family member: A novel approach to teaching breaking bad news. *Studies in Health Technology and Informatics*. 2006;119:67-72.
50. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*. 2000;53:1119-1129.
51. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. *BMJ*. 2015;351:h5651.

52. Sedgwick P. Meta-analyses: heterogeneity and subgroup analysis. *BMJ*. 2013; 346:f4040.
53. Moore PM, Rivera Mercado S, Grez Artigues M, Lawrie TA. Communication skills training for healthcare professionals working with people who have cancer. *The Cochrane Database of Systematic Reviews*. 2013;3:CD003751.
54. Barth J, Lannen P. Efficacy of communication skills training courses in oncology: A systematic review and meta-analysis. *Annals of Oncology*. 2011;22:1030-1040.
55. Otani H, Morita T, Esaki T, et al. Burden on oncologists when communicating the discontinuation of anticancer treatment. *Japanese Journal of Clinical Oncology*. 2011;41:999-1006.

References cited in Appendix 1 only

56. Buckman R. *How to Break Bad News: A Guide for Health Care Professionals*. Baltimore, Md: John Hopkins University Press; 1992.
57. Anandarajah G, Hight E. Spirituality and medical practice: Using the HOPE questions as a practical tool for spiritual assessment. *American Family Physician*. 2001;63:81-89.
58. Bragard I, Razavi D, Marchal S, et al. Teaching communication and stress management skills to junior physicians dealing with cancer patients: A Belgian interuniversity curriculum. *Supportive Care in Cancer*. 2006;14:454-461.

Figure Legends

Figure 1

Flowchart of the study identification process used in a September 2016 systematic review and meta-analysis of the literature on bad or difficult news delivery interventions. The authors updated the search results in February 2017.

^aSome studies were excluded for more than one reason so the listed studies do not equal 53.

^bTwo papers reported on the data collected from a single study and population. The authors treated these as a single study in the analyses and reporting.

Figure 2

Forest plot of the effects of interventions on observer-rated news delivery skill scores from a September 2016 (updated in February 2017) systematic review and meta-analysis of the literature on bad or difficult news delivery interventions. Each solid line represents one study in the meta-analysis, plotted according to the standardized mean difference (SMD). The solid box on each solid line shows the SMD for that study, and the open box on the dashed line represents the pooled SMD. Study IDs given in the following format Author year^{ref}.

Abbreviations: ES indicates standardized mean difference; CI, confidence interval; Bowyer 2010a, Bowyer intervention group 1 (this group watched a 15-minute video on SPIKES); Bowyer 2010b, Bowyer intervention group 2 (this group received a 45-minute didactic lecture on SPIKES, observed an example of bad or difficult news delivery, and participated in a small-group discussion); Bowyer 2010c, Bowyer intervention group 3 (this group received both interventions received by groups 1 and 2); Daetwyler 2010a, Daetwyler intervention group 1 (this group completed an e-learning course on the theory and practice of delivering bad or difficult news); Daetwyler 2010b, Daetwyler intervention group 2 (this group completed the same e-learning course and a simulated news delivery task with feedback).

Figure 3

Forest plot of the effects of interventions on physician confidence in news delivery scores from a September 2016 (updated in February 2017) systematic review and meta-analysis of the literature on bad or difficult new delivery interventions. Each solid line represents one study in the meta-analysis, plotted according to the standardized mean difference (SMD). The solid box on each line solid shows the SMD for that study, and the open box on the dashed line represents the pooled SMD. Study IDs given in the following format Author year^{ref}. Abbreviations: ES indicates standardized mean difference; CI, confidence interval; Bowyer 2010a, Bowyer intervention group 1 (this group watched a 15-minute video on SPIKES); Bowyer 2010b, Bowyer intervention group 2 (this group received a 45-minute didactic lecture on SPIKES, observed an example of bad or difficult news delivery, and participated in a small-group discussion); Bowyer 2010c, Bowyer intervention group 3 (this group received both interventions received by groups 1 and 2).

Figure 1

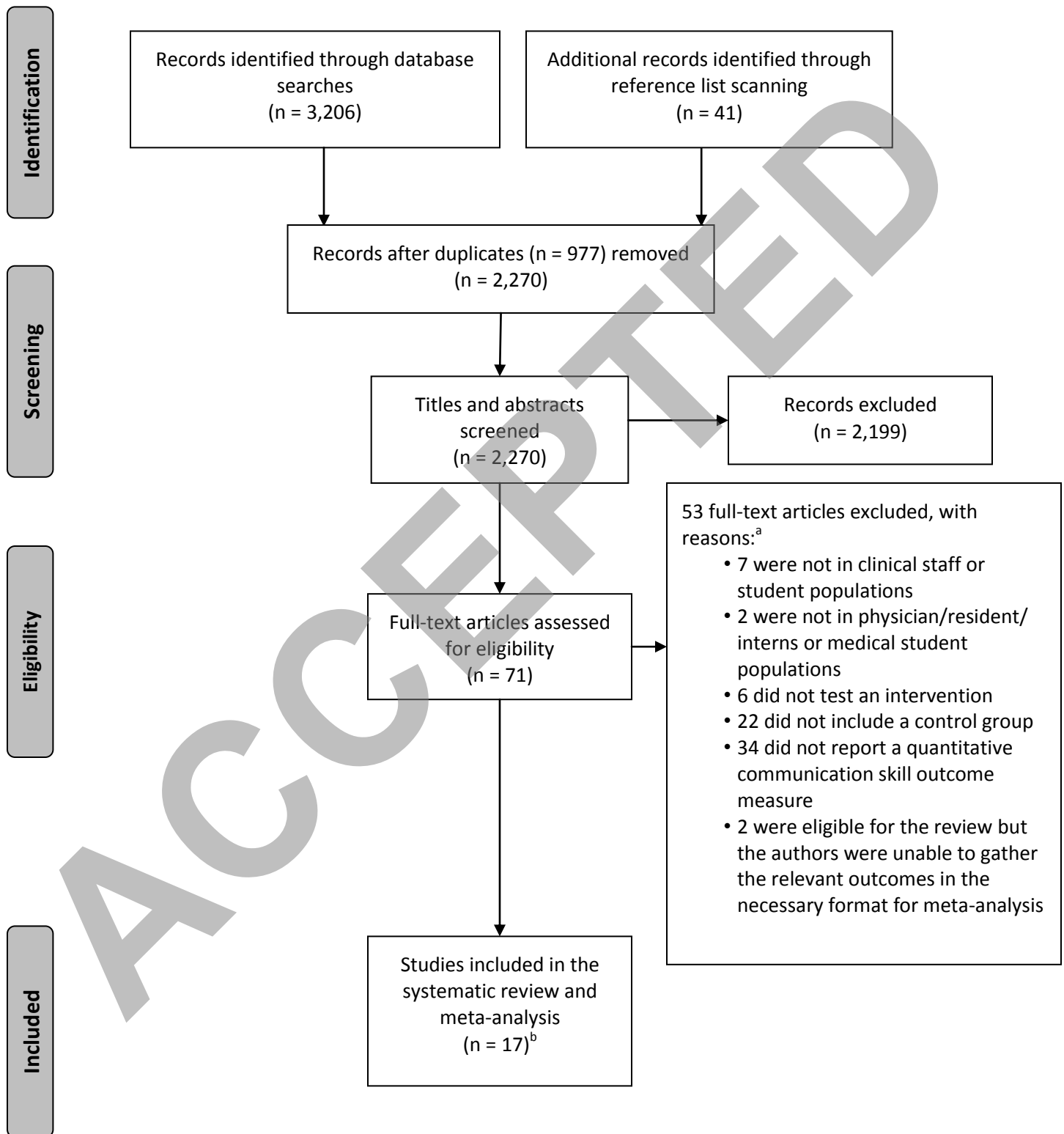


Figure 2

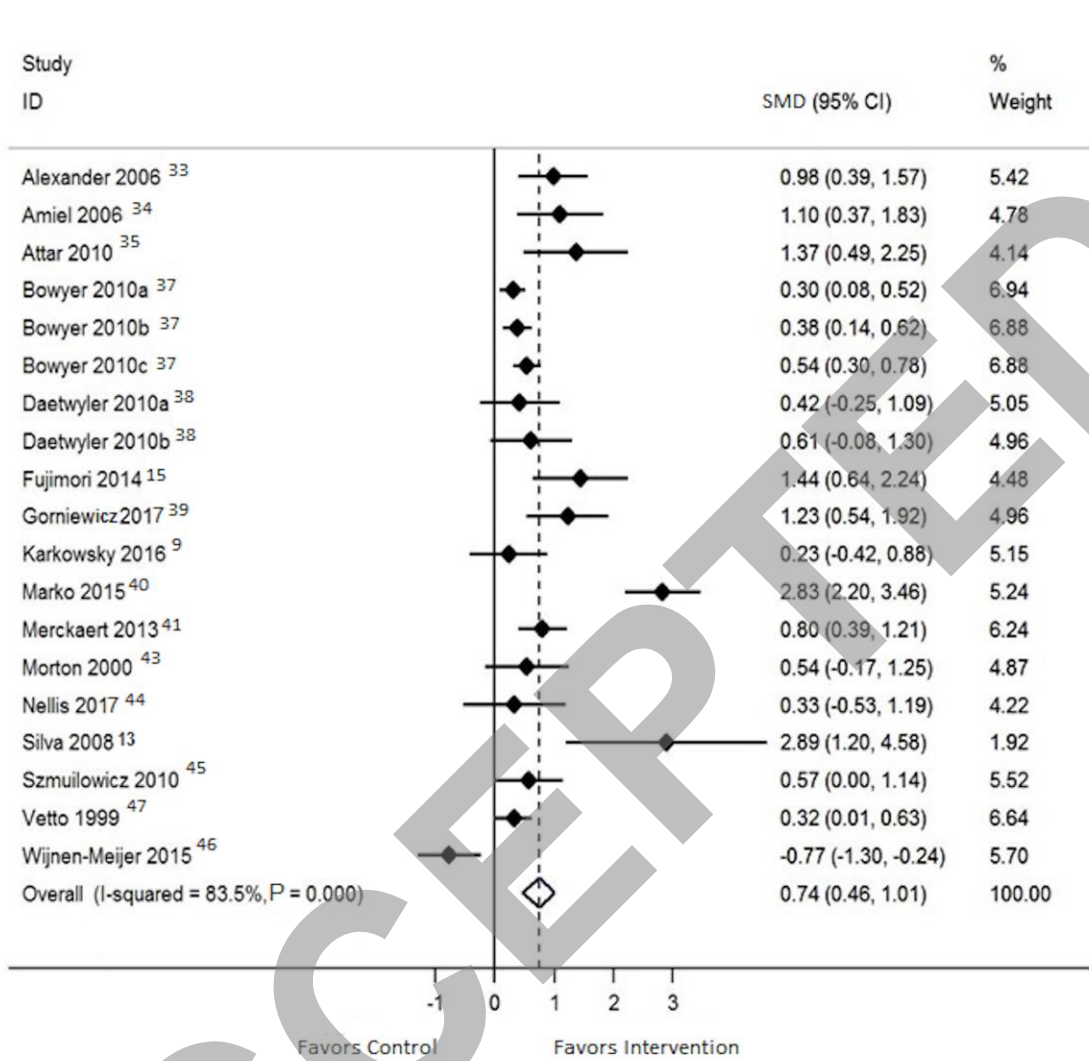
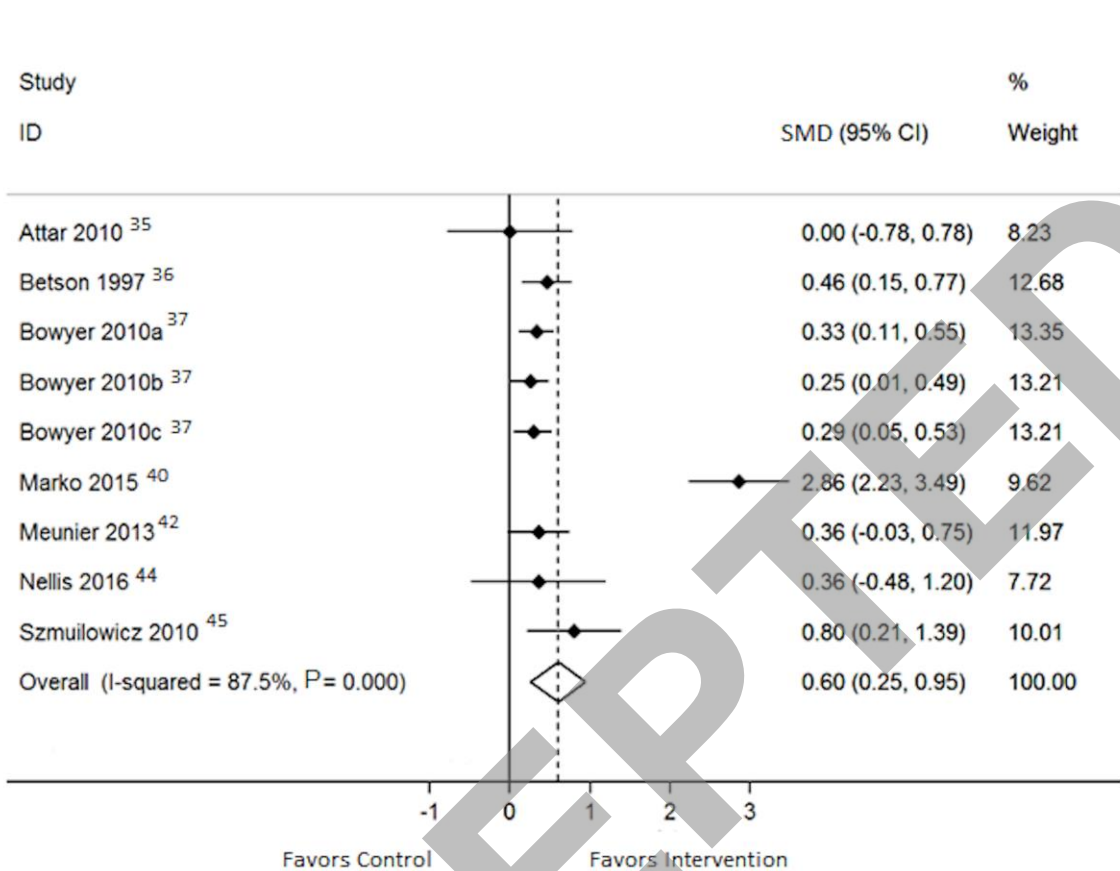


Figure 3



Appendix 1

Participant Demographics and Setting and Design Characteristics of the Studies Included in a September 2016 (Updated in February 2017) Systematic Review and Meta-Analysis of the Literature on Bad or Difficult New Delivery Interventions

First author, year ^{ref}	Participants (recruitment strategy)	Setting (country)	Male, no. (%) ^a	Age (in years), mean ^a	Research design	Intervention	Theoretical basis of the training	Control	Measurement time points
Alexander, 2006 ³³	Medical residents (unclear)	Palliative care (US)	29/57 (52)	NA	CBA	Ninety-minute breaking bad news session delivered as part of a 5-hour communication skills training program that included small-group lectures and discussions, videos, and role-plays. After the 90-minute session, participants undertook 2 hours of communication role-plays.	Unclear	No intervention	CG participants completed one evaluation (timing unclear). IG participants completed the evaluation before and after the intervention (specific time frames not described).
Amiel, 2006 ³⁴	General practitioners (unclear)	Primary care (Israel)	13/34 (38)	IG: 44 CG: 46	CBA	Fourteen 90-minute small-group sessions addressing (1) learning methods of stress management and crisis intervention, (2) coping with emotions when delivering bad or difficult news, (3) communication skills, and (4) role-plays with simulated patients.	Buckman's six-step protocol ⁵⁶	CG participated in a Balint group	All participants completed baseline and post-intervention measures. Specific time frames unclear.

Attar, 2010 ³⁵	Pediatric residents (all residents at the study site were invited to participate in the intervention); controls were students who had not been exposed to the curriculum	Pediatrics (US)	NA	NA	CBA	A training curriculum including three 1-hour sessions on delivering bad or difficult news. Training included didactic teaching and peer-to-peer role-playing. Residents who could not attend the training viewed materials via a web platform.	Integrated three frameworks: Buckman's six-step protocol, ⁵⁶ SPIKES, ¹⁴ and HOPE model for spirituality ⁵⁷	No intervention	CG participants completed one assessment during their first year as subspecialty pediatric fellows (specific time point unclear). IG participants completed assessments during the first year of their residency and at the end of their third year (specific time point in the first year unclear).
Batson, 1997 ³⁶	Medical students (all students in their third year were invited)	University setting (Hong Kong)	NA	NA	RCT	Participants received a 3-hour seminar on breaking bad or difficult news as part of a wider communication training course. During this, IG participants watched a culturally appropriate (Cantonese) news delivery video.	Unclear	CG also received the 3-hour seminar on breaking news, but they watched a non-culturally appropriate (British) video.	Measured at four time points: baseline, immediately after viewing the respective videos, between 3–6 weeks after viewing the videos, and 4 months after viewing the videos.
Bowyer, 2010 ³⁷	Medical students (unclear, but	University setting (US)	NA	NA	RCT	There were three intervention levels: (1) a 15-minute video on	SPIKES ¹⁴	No intervention	News delivery skill was measured once after the

	it seems all students were entered into the study as part of their training)					SPIKES; (2) a 45-minute didactic lecture on SPIKES, observation of a faculty facilitator showing a good example of news delivery, and small-group discussion; and (3) a 15-minute video and 45-minute didactic lecture on SPIKES, observation of a faculty facilitator showing a good example of news delivery, and small-group discussion.			intervention. Student sense of preparedness to break news was measured before the news delivery task, immediately after the task, and at the end of the 12-week rotation.
Daetwyler, 2010 ³⁸	Medical interns (participants volunteered from a pool of 62 interns)	Oncology (US)	40/55 (73)	29	nRCT (participants were allowed to switch their conditions if their allocation was inconvenient)	There were two intervention levels: (1) an e-learning course on the theory and practice of delivering bad or difficult news and two videos demonstrating communication skills and (2) the e-learning course as well as a simulated news delivery task with feedback.	Unclear	No intervention	All participants completed a baseline (3 weeks prior to the intervention) and post- (7–8 weeks after baseline) assessment.
Fujimori, 2014 ¹⁵	Oncologists (recruited from two hospitals)	Oncology (Japan)	26/30 (87)	39	RCT	Two-day long training course involving group discussion, didactic teaching with text and videos, role-playing, and	SHARE ¹⁵	No intervention	Baseline and after the intervention (with a 2-week gap between the assessments).

						discussions on the role-playing.			
Gorniewicz, 2017 ³⁹	Family medicine and internal medicine residents (participation was part of their usual training)	Unclear, but seems it was held in a university setting (US)	20/38 (53)	29	RCT	A 60-minute e-learning module delivered as part of a wider communication skills training course. The module included clips of interviews with patients and quiz questions.	No single theoretical model, but was partially informed by SPIKES ¹⁴	No intervention	Baseline and 2 weeks later (after the IG had received the intervention).
Karkowsky, 2016 ⁹	House staff in obstetrics and gynecology (unclear)	Obstetrics (US)	IG: 4/17 (22) CG: 3/18 (18)	NA	RCT	A 10-minute simulation session that covered similar material to that viewed by participants in the CG but that also reviewed the participant's baseline simulation performance.	Buckman's six-step protocol ⁵⁶	CG received a 10-minute lecture covering breaking bad news principles and skills but did not review baseline performance.	Baseline and post- (between 2–12 weeks after the baseline) assessment. There was also a 6-month follow up, but by this point, all participants had had undergone both conditions.
Marko, 2015 ⁴⁰	Medical students (recruited voluntarily over a 6-month period via convenience sampling)	Obstetrics (US)	NA	NA	CBA	One hour of testing and 2 hours of training involving didactic teaching, observation of faculty demonstrating good practice of news delivery, role-play with faculty feedback, and provision of course	SPIKES ¹⁴	One hour of testing and the traditional medical curriculum (consisting of clinical apprenticeshi	Baseline and post-assessment (with between 4–5 weeks between assessments).

						materials.		p and observation of preceptors counseling patients with early pregnancy loss).	
Merckaert, 2013 ^{41b}	Medical residents (recruited via letters and phone calls to Belgian French-speaking hospitals)	Oncology (Belgium)	IG: 16/48 (33) CG: 19/47 (40)	28	RCT	A 30-hour communication skills and stress management training course covering dyadic and triadic consultations. Training involved teaching on bad or difficult news delivery and role-playing with facilitator feedback.	BIC-CST ⁵⁸	Wait list	Assessments were made before randomization for both groups, after completion of the training program for the IG, and 8 months after the first assessment for the CG.
Meunier, 2013 ^{42b}	Medical residents (recruited via letters and phone calls to Belgian French-speaking hospitals)	Oncology (Belgium)	IG: 16/50 (32) CG: 19/48 (40)	28	RCT	A 30-hour communication skills and stress management training course covering dyadic and triadic consultations. Training involved teaching on bad or difficult news delivery and role-playing with facilitator feedback.	BIC-CST ⁵⁸	Wait list	Assessments were made before randomization for both groups, after completion of the training program for the IG, and 8 months after the first assessment for the CG.
Morton, 2000 ⁴³	Intensive care	Intensive care (UK)	IG: 14/32 CG: 40 ^c	IG: 38 CG: 40 ^c	RCT	Interactive workshop with short presentations,	Unclear	No intervention	Baseline (3–4 weeks before the

	physicians (unclear)		(44) CG: 16/32 (50) ^c				exercises, discussion, and role-plays. Workshops included mixed groups of 12–16 doctors and nurses.		intervention), post- (4–6 weeks after the intervention), and a 6-month follow-up assessment.
Nellis, 2017 ⁴⁴	Medical residents (all second-year residents based on the pediatric intensive care unit where the study took place were invited to participate over a 2-year period)	Pediatric intensive care (US)	NA	NA	CBA	A 40-minute simulation- based training workshop. Participants worked in pairs on a simulation of caring for a critically ill child. One resident discussed the neurologic implications of a prolonged arrest with the parent, and the second resident informed the parent that the child had died after a second arrest. Residents received feedback on their communication skills from a trained faculty facilitator.	Unclear	CG received informational resources focused on communicati ng with children and their families at the end of life.	Self-perceived competency was measured before and after the intervention. Observer-rated communication skills were measured once (6 weeks after the intervention).
Silva, 2008 ¹³	Pediatric medical residents (participated as part of their required teaching activities)	Pediatrics (Puerto Rico)	NA	NA	CBA	Four 1-hour sessions, each involving didactic teaching, viewing a good example of bad or difficult news delivery, viewing videos of news delivered both appropriately and	SPIKES ¹⁴	No intervention	Measures were taken 3 weeks after the intervention was completed.

						inappropriately and rating these using a checklist, and role-plays with feedback.			
Szmuilowicz, 2010 ⁴⁵	Internal medicine residents (all second-year residents at the hospital were invited by e-mail, with a follow-up letter at 2 weeks)	Palliative care (US)	IG: 11/21 (52) CG: 20/28 (71)	IG: 29 CG: 30	nRCT (two participants switched to the CG after they were randomized to the IG)	Six hours of training involving small-group discussions, modeling of conversations, didactic teaching, and role-plays with feedback from trained faculty.	Unclear	No intervention	Baseline and after the intervention. The average amount of time between the intervention and the second assessment was 14 weeks.
Vetto, 1999 ⁴⁷	Medical students (unclear, but students were recruited from a university)	Unclear (US)	IG: 32/69 (46) CG: 45/86 (52)	NA	CBA	Delivered as part of a new curriculum for medical students. It included 6 hours of panel discussions with clinicians, group discussions, and role-playing.	Buckman's six-step protocol ⁵⁶	Old medical curriculum	Assessment was conducted at one time point following the intervention during students' third-year family practice clerkships.
Wijnen-Meijer, 2015 ⁴⁶	Medical students (recruited on a voluntary basis via ads)	Hospital outpatient department (Germany and The Netherlands)	IG: 10/30 (33) CG: 7/29 (24)	NA	CBA	A new medical curriculum featuring early clinical experience, integration of science and practice, and progressive increase of clinical responsibility. The curriculum	Unclear	Old medical curriculum	Assessment was conducted at one time point following the intervention (for the IG this was 2 weeks prior to their graduation and for

allocated more time to training in communication skills, although specific details of training in bad or difficult news delivery were not provided.

the CG this was 3 months prior to their graduation).

Abbreviations: US indicates United States; NA, not available; CBA, controlled before-after study; CG, control group; IG, intervention group; SPIKES, six-step framework for bad or difficult news delivery (steps are as follows: setting up the interview, assessing the patient's perception of the situation, obtaining the patient's invitation to deliver the news, giving knowledge and information to the patient, addressing the patient's emotions empathically, and providing a summary and discussing prognosis and treatment options); HOPE, model of spirituality which instructs physicians to elicit information from their patients concerning their sources of hope, organized religion, personal spirituality, and the effects of this on their medical care; RCT, randomized controlled trial; nRCT, non-randomized controlled trial; SHARE, four-step protocol for delivering bad or difficult news (steps are as follows: create a supportive environment, consider how to deliver the news, discuss additional information that patients would like to know, and provide reassurance and emotional support); BIC-CST, Belgian Interuniversity Curriculum - Communication Skills Training; UK, United Kingdom; EDHEP, European Donor Hospital Education Programme.

^aWhere available, the authors report the overall sample number and percentage or mean but some studies only provided these by IG or CG.

^bThese papers report data collected from a single study and population. The authors treated these as a single study in the analyses and reporting. That is, only the data reported by Merckaert and colleagues⁴⁰ was included in the main meta-analysis, and only the data reported by Meunier and colleagues⁴¹ was reported in the secondary meta-analysis.

^cThis data is based on the overall sample included in the study, which included 32/64 (50%) nurses. Nurse outcome data is not included in the meta-analyses.