Implementation Science

Facilitating Implementation of Research Evidence (FIRE): an international cluster randomised controlled trial to evaluate two models of facilitation informed by the Promoting Action in Research Implementation in Health Services (PARIHS) framework.

--Manuscript Draft--

Manuscript Number:	IMPS-D-18-00228R2
Full Title:	Facilitating Implementation of Research Evidence (FIRE): an international cluster randomised controlled trial to evaluate two models of facilitation informed by the Promoting Action in Research Implementation in Health Services (PARIHS) framework.
Article Type:	Research
Funding Information:	EUFP7 (FP7/2007-2013 under grant agreement n° 223646.) — Prof. Kate Seers FP7 Health (n° 223646.) — Prof. Kate Seers
Abstract:	Background Health care practice needs to be underpinned by high quality research evidence, so that the best possible care can be delivered. However, evidence from research is not always utilised in practice. This study used the Promoting Action of Research Implementation in Health Services (PARIHS) framework as its theoretical underpinning to test whether two different approaches to facilitating implementation could affect the use of research evidence in practice. Methods A pragmatic clustered randomised controlled trial with embedded process and economic evaluation was used. The study took place in four European countries across 24 long term nursing care sites, for people aged 60 years or more with documented urinary incontinence. In each country, sites were randomly allocated to standard dissemination, or one of two different types of facilitation. The primary outcome was the documented percentage compliance with the continence recommendations, assessed at baseline, then at 6, 12, 18 and 24 months after the intervention. Data were analysed using STATA10, linear regression models were fitted to scores for compliance with continence recommendations, adjusting for clustering. Results Quantitative data were obtained from reviews of 2313 records. There were no significant differences in the primary outcome (documented compliance with continence recommendations) between study arms and all study arms improved over time. Conclusions This was the first cross European randomised controlled trial with embedded process evaluation that sought to test different methods of facilitation. There were no statistically significant differences in compliance with continence recommendations between the groups. It was not possible to identify whether different types and "doses" of facilitation were influential within very diverse contextual conditions. The process evaluation (linked paper) revealed the models of facilitation used were limited in their ability to overcome the influence of contextual factors.
Corresponding Author:	Kate Seers, PhD DSc University of Warwick Warwick Medical School Coventry, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Warwick Warwick Medical School
Corresponding Author's Secondary Institution:	
First Author:	Kate Seers

First Author Secondary Information:	
Order of Authors:	Kate Seers
	Jo Rycroft-Malone
	Karen Cox
	Nicola Crichton
	Rhiannon Tudor Edwards
	Ann Catrine Eldh
	Carole Estabrooks
	Gill Harvey
	Claire Hawkes
	Carys Jones
	Alison Kitson
	Brendan McCormack
	Christel McCullan
	Carole Mockford
	Theo Niessen
	Paul Slater
	Angie Titchen
	Teatske van der Zijpp
	Lars Wallin
Order of Authors Secondary Information:	
Response to Reviewers:	Reviewer 1 Reviewer Comment a)If ACT was collected at baseline, I don't follow how it is listed as a secondary 'outcome'. I guess if only measured at baseline it should be reported much earlier in the results section. Response:ACT was also collected at 12 and 24 months, however, this data is not available for all sites at 12 months and 24 months. In this paper it is the baseline data (the organisational context at the point of implementation) that seemed most relevant and we are really using ACT as an explanatory variable rather than as a secondary outcome variable in this paper. We would prefer to leave the results related to ACT at the end of the findings section. Comment: b)Typically, ICCs are higher for process measures than for more distal outcomes. process measures in primary care often have an ICC of 0.1 and outcome measures more like the stated 0.01. I appreciate the additions made in regard to ICCs - I might suggest that a revised sample size calculation could be helpful for readers. Response: Retrospective sample size calculations are a statistically controversial issue and we do not think it will help the reader in this case. We have now included, at the suggestion of one of the other referees, the post estimation ICCs that follow from the fitting of the regression models. Comment: c)If ACT was similar across sites, how do we understand differences observed across countries and sites? maybe ACT doesn't capture contextual factors relevant to improvement in these processes or relevant to responsiveness to the intervention? Response: Thank you for this comment. We address this in the discussion section of our paper. The issue of whether ACT is sensitive to change we do not discuss as we have only used baseline data from ACT in this paper so do not examine change in ACT scores. The team of researchers that developed ACT are engaged in studies that are examining the sensitivity of ACT to change and this may be reported in a subsequent paper, but as stated in our response to your first point, we do not have data for all sites at 12 and 24 months.

Reviewer 3
Reviewer Comment
a)The primary outcomes are three measures of 'compliance with continence recommendations' (resident screened; assessed; treatment plan; specialist referral) (abstract and in the methods p11). We are not clear how these are measured. Is this a percentage compliance at the level of the patient or the cluster? (nursing care site). If the latter, the regression models (tables 2 4 6 8) only include 24 units. We have assumed in the following comments that the analysis is based on patient-level outcomes. It would be helpful to expand the description of the outcomes in the methods section to explain how the measures have been constructed. Add a N= at the bottom of all the tables, to give the overall population, making clear whether these are sites or patients.
Response: Percentage compliance is at the resident level. A sentence has been added on page 9 in the outcome measures section. Supplementary file 1 provides details of the components of each of the recommendations. For each resident the percentage compliance with a recommendation is the total number of components of that recommendation that the documentation indicates have been complied with divided by the number of components and expressed as a percentage.

N=2313 has been added to the bottom of tables 1-7 to make clear that these tables present an analysis at the resident level.

Table 8 presents a summary of the staff responses to the ACT questionnaire, N=725 has been added to the title of this table so it is clear how many staff responses the table is based on overall.
Comment b)The analysis of the primary outcomes (three measures of compliance with continence recommendations) uses linear regression models. This is a cluster randomised trial with multiple levels (country, site, staff, patient). In a cluster randomised trial it is usual to perform an analysis which adjusts for clustering at the unit of randomisation, which in this case is nursing care site, and this is what is described. 'Regression' covers a range of different analyses, and in this case presumably some sort of multilevel model was used to adjust the standard errors to take site level clustering into account so it would be helpful to add in exactly which stata command was used to fit the regression models (e.g. mixed, xtmixed). It would be helpful to add a statement that the assumptions of linear regression have been examined and the data meets those assumptions,
Response: We agree with the referee that this study could be considered to have multiple levels. We have considered data at the resident level and clustering is at the site level. Country has been included as a covariate. With regard to the primary outcomes we have no information about staff so this cannot be considered as a level in the trial design.

In the previous version of the paper the models were linear regression models in which the standard errors were specified as robust (cluster) with the site as the cluster variable. This corrects the standard errors through the sandwich method (Huber-White method), inflating the se's to correct for the clustering.

We have also fitted a multi-level mixed effect linear regression model (using the STATA15 mixed command, again with site as the level variable and SE set to robust(cluster)). The estimates are very similar to those from the simpler model we reported in the previous version, some of the se's are increased a little and none of the conclusions from the models are changed. We have decided to update the results tables 2, 4, 6 to show the results from the multi-level mixed effect linear regression to be sure we have taken full account of the clustering. We have amended the explanation of the model fitting on page 11, to clarify what was done.

We have added a statement about the assumptions on page 13-14.
Comment
c)ICCs were calculated on the baseline data and we find this confusing. Once a multilevel model has been fitted to the data, it is usually possible to extract an overall ICC at the level of clustering which would be based on all data points (not just baseline) - see https://www.stata.com/features/overview/intraclass-correlations-for-multilevel-models/

Response:
The reason for calculating the ICCs at baseline is these are the values which, had they been available, we would have used in the sample size calculation.

We have, as you suggested, calculated the post-estimation ICC for each of the regression models. These post estimation ICCs are reported at the end of tables 2, 4 and 6 and in the text relating to the results for each of the three compliance variables (pages 14-16).

This analysis was not available in version 10 of STATA (the version we had been using for the analysis). We have therefore rerun all the analyses in STATA version 15 and throughout the text updated STATA10 to STATA15.
Comment: d)There were only 24 sites (clusters) and both 'country' and 'intervention' are fitted at a site level and have a total of 5 individual covariates between them - it is doubtful there is enough data here to provide stable estimates: We'd recommend that the authors try removing 'country' from the analysis as a sensitivity analysis to check that results are similar.
Response: We agree with the reviewers that the number of sites is small relative to the number of levels of the covariates. As the reviewers suggest have re-run the analysis removing country from the analysis as a sensitivity check. There is still no significant intervention effect for any of the three recommendations if country is removed. The results are similar, so the estimates appear to be reasonably stable. This has been noted in on page 13-14.
Comment e)Patients were recruited to this trial post randomisation - which means that recruiters already knew which arm of the trial a site was in before consenting patients. This lack of allocation concealment can lead to differential recruitment in a cluster randomised trial - e.g. differences in numbers recruited or the type of patients recruited. We can see little information on which to judge whether or not this was a problem. We can't find a consort flow diagram and this would be helpful (and recommended by CONSORT) - how many patients were approached but did not consent, and how many consented but did not provide outcome data? Baseline compliance in the control group was much higher than in the intervention group (table 3) and this could be an indication of differential recruitment. Perhaps the authors can make some comment on this.
Response: On page 10 in the section Sample size and power calculation it says “Consent was sought at cluster and at individual level, the former before randomisation and the later after randomisation.” On page 11 in the section on allocation concealment and blinding it said “It was not possible to blind sites to intervention, although research fellows who collected data were initially blinded to intervention group.”
This sentence has been reworded (on page 11) to make it clear that where it was necessary to obtain consent from individual residents for outcome data collection this consent was obtained by the research fellow who was initially blind to the intervention allocated for the site.

To clarify for the reviewers, consent from residents was not necessary for access to records in Netherlands, Sweden or Ireland. All data from records was collected by the country research fellow who was blinded to the intervention group to which the care home had been allocated, so recruitment of resident records would not have been influenced by lack of allocation concealment. In UK, consent from residents (or their family) was necessary for access to the resident record. This consent was obtained by the research fellow who was unaware of which intervention the care home had been allocated. We noted that once the research fellows were in the long term care setting, the blinding could inadvertently be broken by the site, for example, mentioning the name of an external facilitator working with them, and thus revealing the allocation of that site. In all countries the consent of the resident or their family was necessary from the collection of EQ5D, this consent was obtained by the research fellow.
The study flow diagram is in supplementary file 4.

Comment f)How similar were patient demographics across groups? The paper would benefit from a table of baseline characteristics, by group. Provide mean(SD) for continuous variables, number(%) for categorical and median (IQR) for ordinal or non-normal, and include both the number included for each measure plus the overall total. It is not good practice to compare the groups using statistical tests (as described on

	page 11 and shown in table 1). Response: The only resident demographic information that was collected was age and gender. We prefer not to introduce a further table to report this information by group, but we have included it within the text on page 12 in the description of resident sample. Whilst we agree with the reviewers that using statistical tests to compare groups is not considered by everyone to be good practice, however it is a widely used practice. We only do this for demographic data and some secondary outcomes because we think some readers would expect (or prefer) to have evidence of statistical tests alongside statements of similarity or difference. With regard to Table 1 we have included 95% CI's for those who prefer this approach to describing similarity or difference between groups. Comment: g)EQ5D-VAS is reported (table 1) at 24 months only (because of least amount of missingness at that time point). It was collected at several time points, suggesting it might be an outcome measure, but it is not reported as such. Response: The reviewers observation that EQ5D might have been a secondary outcome measure is correct. However, there were difficulties with the process of getting resident consent for collection of this data and process of getting this measure completed with the resident. Over time the issues related to consent and process were resolved sufficiently so that at the 24 month data collection point this measure was available for 77% of residents. Given the difficulties with collecting this data earlier in the trial we decide this measure could not be used as a secondary outcome. We did however, feel that the more complete 24 month data was helpful in providing a description of the health status of the residents. Comment: h)Table 3 and 5 both suggest an effect in favour of the 2 intervention arms, and the effect is quite close to the 15 percentage point difference anticipated in the power calculation (Table 7 does not). This, combined with the very wide confidence intervals, might indicate that there is a hint of an effect here, but that there is insufficient power to detect a difference. We like the way the authors have used the findings from the process evaluation to explain the results: however, they cannot write off the idea that there may have been an effect, and the lack of power may be more of an issue than suggested in the conclusion. Response: We agree with the reviewers' observation that given the lack of power we cannot write off the idea that there may have been an effect from one or both of the intervention arms, and this is addressed in the limitations. We feel that the text on page 14 already identifies that Tables 3 and 5 show some improvement in the intervention arms, but given the large variability associated with the means reported and the lack of power we are reluctant to put any further emphasis on these results. As the reviewers note the confidence intervals are very wide. Comment: i)Secondary outcomes at patient level seem to have been reported as a difference between baseline and 24 months, rather than compared by group. These outcomes should be reported as described in the methods, using Anova or chi squared tests, as appropriate, or even using regression methods Response: All the secondary outcomes at patient level are binary outcomes we have therefore used chi-squared test to compare the groups at 24 months and this information has been added on page 16-17. As further exploration of the differences found between baseline and 24 months we feel it is helpful to look at changes within the intervention groups, so the changes within intervention group previously reported remain in the text. Comment: j)Given the lack of power in the study, we would very tentatively suggest that the authors give consideration to combining both arms and comparing facilitation against control, although this analysis was not anticipated and would have to be reported very carefully. Response: Although the study was not designed with the intention of combining facilitation arms, the analysis has been rerun comparing facilitation (combining the groups Type A and type B) against control and there is no evidence that the facilitation is more effective than control.
Additional Information:	
Question	**Response**
<b>Is this study a clinical	Yes

trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'.	
We require registration of all clinical trials that are reported in manuscripts submitted to the journal. More information about trial registration, including the trial registries that currently meet all of the ICMJE guidelines, can be found in the FAQ section of "About ICMJE" at http://www.icmje.org/about-icmje/faqs/clinical-trials-registration/.<p>Please provide the following information where prompted:<hr>Enter the Trial Registration Number: as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	ISRCTN11598502
Enter the name of the registry: as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	Current Controlled Trials
Enter the URL of the trial registry record: as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	https://doi.org/10.1186/ISRCTN11598502
Enter the date that you registered your trial (in mm/dd/yyyy format): as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	04-02-2010
Enter the date of enrolment of the first participant to the trial (in mm/dd/yyyy	01-03-2010

format): as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	
Was your trial registered before the first participant was enrolled? (i.e. prospectively registered) as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	Yes
Within your manuscript, have you also included details of your trial registration at the end of your abstract? • Name of the registry • Trial registration number • Date of registration • URL of trial registry record *Example: Trial registration: ISRCTN, ISRCTN12345678. Registered 28 September 2014, http://www.isrctn.com/ISRCTN12345678* as follow-up to "Was your trial registered before the first participant was enrolled? (i.e. prospectively registered)"	I confirm I have provided trial registration details at the end of the abstract

Facilitating Implementation of Research Evidence (FIRE): an international cluster randomised controlled trial to evaluate two models of facilitation informed by the Promoting Action in Research Implementation in Health Services (PARIHS) framework.

Full list of authors, Institutional addresses and emails

Kate Seers*, Warwick Research in Nursing, Warwick Medical School, University of Warwick, UK. Kate.seers@warwick.ac.uk

*Corresponding Author

Jo Rycroft-Malone, School of HealthCare Sciences, Bangor University, UK. j.rycroft-malone@bangor.ac.uk

Karen Cox, Fontys School of People and Health Studies, Fontys University of Applied Sciences, Eindhoven, The Netherlands k.cox@fontys.nl

Nicola Crichton, School of Health and Social Care, London South Bank University, 103 Borough Road, London SE1 0AA, UK. crichtnj@lsbu.ac.uk

Rhiannon Tudor Edwards, Centre for Health Economics and Medicines Evaluation (CHEME) Bangor University, UK. r.t.edwards@bangor.ac.uk

Ann Catrine Eldh, Linkoping University, Faculty of Medicine and Health Science, Department of Nursing, Linkoping, Sweden and Department of Neurobiology, Care Sciences and Society, Division of Nursing, Karolinska Institutet, Stockholm, Sweden. Ann.catrine.eldh@liu.se

Carole A Estabrooks, Faculty of Nursing, University of Alberta, Edmonton, Alberta, Canada carole.estabrooks@ualberta.ca

Gill Harvey, Adelaide Nursing School, University of Adelaide, Australia.
gillian.harvey@adelaide.edu.au

Claire Hawkes Senior Research Fellow, Clinical Trials Unit, Warwick Medical School, University of Warwick, UK. C.A.Hawkes@warwick.ac.uk

Carys Jones, Centre for Health Economics and Medicines Evaluation (CHEME), Bangor University, UK.
CHEME, Bangor University, UK. c.l.jones@bangor.ac.uk

Alison Kitson, College of Nursing and Health Sciences, Flinders University, South Australia, Australia.
alison.kitson@flinders.edu.au

Brendan McCormack, Division of Nursing, Queen Margaret University Edinburgh, UK
bmccormack@qmu.ac.uk

Christel McMullan Research Fellow, Institute of Applied Health Research, University of Birmingham,
UK, c.mcmullan@bham.ac.uk

Carole Mockford, Warwick Medical School, University of Warwick, UK.
Carole.mockford@warwick.ac.uk

Theo Niessen, Fontys School of People and Health Studies, Fontys University of Applied Sciences,
Eindhoven, The Netherlands t.niessen@fontys.nl

Paul Slater Institute of Nursing and Health Research, Ulster University, Shore Rd Belfast. Northern
Ireland. pf.slater@ulster.ac.uk

Angie Titchen, Institute of Nursing and Health Research, Ulster University, Jordanstown, UK,
a.titchen@ulster.ac.uk

Teatske van der Zijpp, Fontys School of People and Health Studies,

Fontys University of Applied Sciences, Eindhoven, The Netherlands, T.vanderzijpp@fontys.nl

Lars Wallin, School of Education, Health and Social Studies, Dalarna University, Falun, Sweden.

Department of Neurobiology, Care Sciences and Society, Division of Nursing, Karolinska Institutet,

Stockholm, Sweden and Department of Health and Care Sciences, The Sahlgrenska Academy,

University of Gothenburg, Gothenburg, Sweden. lwa@du.se

Facilitating Implementation of Research Evidence (FIRE): an international cluster randomised controlled trial to evaluate two models of facilitation informed by the Promoting Action in Research Implementation in Health Services (PARIHS) framework

Abstract

Background

Health care practice needs to be underpinned by high quality research evidence, so that the best possible care can be delivered. However, evidence from research is not always utilised in practice. This study used the Promoting Action of Research Implementation in Health Services (PARIHS) framework as its theoretical underpinning to test whether two different approaches to facilitating implementation could affect the use of research evidence in practice.

Methods

A pragmatic clustered randomised controlled trial with embedded process and economic evaluation was used. The study took place in four European countries across 24 long term nursing care sites, for people aged 60 years or more with documented urinary incontinence. In each country, sites were randomly allocated to standard dissemination, or one of two different types of facilitation. The primary outcome was the documented percentage compliance with the continence recommendations, assessed at baseline, then at 6, 12, 18 and 24 months after the intervention.

Data were analysed using STATA15, multi-level mixed effects linear regression models were fitted to scores for compliance with the continence recommendations, adjusting for clustering.

Results

Quantitative data were obtained from reviews of 2313 records. There were no significant differences in the primary outcome (documented compliance with continence recommendations) between study arms and all study arms improved over time.

Conclusions

This was the first cross European randomised controlled trial with embedded process evaluation that sought to test different methods of facilitation. There were no statistically significant differences in compliance with continence recommendations between the groups. It was not possible to identify whether different types and "doses" of facilitation were influential within very diverse contextual conditions. The process evaluation (linked paper[36]) revealed the models of facilitation used were limited in their ability to overcome the influence of contextual factors.

Trial Registration: Current Controlled Trials ISRCTN11598502. Date 4/2/10.

https://doi.org/10.1186/ISRCTN11598502

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 223646.

Keywords (up to 10):

Facilitation, implementation, PARIHS, urinary incontinence, context, older people, RCT

Background

It is important that health care practice is underpinned by high quality research evidence, so that the best possible care can be delivered. However, evidence from research is not always utilised in practice.[1-3] This study used the Promoting Action of Research Implementation in Health Services (PARIHS) framework[4] as its theoretical underpinning to test whether two different approaches to facilitation could affect the use of research evidence in practice. The PARIHS framework was built upon an argument that three factors influence the uptake of research evidence in practice: the nature (strength) of the evidence, the context in which it is used, and the extent of facilitation (or help) that people have to use the evidence. The published protocol for this study[5] and an online summary report for the funder[6] contains further details.

Consistent with recent calls for an increase in theory-based implementation research,[7] we used the PARIHS framework and identified two alternative types of facilitation to evaluate within the FIRE Study. We chose to evaluate facilitation because whilst it is a promising approach to implementation; it has received relatively little attention and the limited results available of its effectiveness were mixed.[8-10] Facilitation has been described as a process and a role.[11] More recently it has been argued[12] "conceptual ambiguities" challenge our understanding of facilitation's effectiveness and we do not know how to "appropriately set the degree of facilitation." It is clear from the literature that the role and effectiveness of facilitation in implementing evidence into practice needs to be explored and tested. This study was novel in scale with a cross-country setting, and in that it sought to compare facilitation approaches that varied in terms of focus, duration and intensity.

Urinary incontinence in long term care settings is a major issue and was thus selected as an exemplar for evaluating different approaches to implementing evidence into practice. Incontinence is a "discrediting and stigmatising" condition that affects quality of life.[13] It has a high prevalence in long term care settings, between 40-70%,[14] and it is a key priority within international health

policy.[15] The relevance and fit of the PARIHS framework in long term care settings for older people[16] highlighted that the factors discussed as important for change in their setting showed a good fit with those identified in the PARIHS framework, and recommended its use in these settings. We designed the FIRE trial to test two different approaches to facilitation and compare these against standard dissemination of recommendations for continence promotion.[5]

Aims: We aimed to extend knowledge of facilitation as a process for getting research evidence into practice by testing the effectiveness of and evaluating the contribution two different models of facilitation can make to implementing evidence based urinary continence recommendations into practice.

The objectives of the study were to:

1) Extend existing knowledge of facilitation as a process for translating research evidence into practice.

2) Evaluate the feasibility and effectiveness of two different models of facilitation in promoting the uptake of research-based recommendations on continence promotion, compared with standard dissemination.

3) Advance existing knowledge of guideline implementation in healthcare, with a particular focus on understanding the impact of contextual factors on the processes and outcomes of implementation.

4) Implement a pro-active dissemination strategy that complements the design of the study and facilitates the diffusion of the study findings to a wide policy and practice community throughout Europe and beyond. This objective is not considered further in this paper.

Methods

Design

A pragmatic cluster randomised controlled trial with embedded process and economic evaluation was undertaken. The process evaluation is reported in a linked paper (Rycroft-Malone et al[36]).

Participants

Staff: An Internal facilitator (a member of staff from the long-term care setting) nominated in each intervention site to work with external facilitators (EFs) to implement the urinary incontinence (UI) recommendations.

Residents: aged 60 years or more with documented urinary incontinence.

Setting

The study took place in four European countries (England, Sweden, Netherlands, Republic of Ireland), and each country planned to recruit six long term nursing care sites (nursing homes and other residential settings with long term nursing care) (total 24 sites) for people aged 60 years or more with documented urinary incontinence. All settings had publicly funded places.

The intervention

In arm one, the eight settings randomised to the standard dissemination control group had the urinary continence recommendations and a PowerPoint presentation on implementation (based on one utilised by Rycroft-Malone et al[17]) sent to the head of each site. Both the intervention groups also received the same as the standard dissemination sites.

In addition, EFs prepared two different facilitator development programmes, each of which involved an initial residential programme, followed by virtual support (monthly telephone group supervision and email communication) for the internal facilitators (IFs) in implementing the UI recommendations. Arm two received a type of facilitation that we termed 'type A', which is a goal focused approach to facilitation based on principles of quality improvement, management studies

and organisational learning. This involved a 12 month programme for IFs nominated by each of the eight sites in this arm. This started with the IFs taking part in a three day residential programme run by two EFs (GH & AK), followed by 10 days over 12 months to work locally on the implementation and evaluation of recommendations, supported by 12 half days for monthly teleconferences and self-directed study (16 days in total).

Arm three received a type of facilitation that we termed 'type B', which is underpinned by principles of stakeholder empowerment and overcoming external and internal obstacles to using research evidence in practice. This is achieved through the creation of workplace cultures of effectiveness in which work-based learning as inquiry is valued and supported at all levels of the organisation. This approach is informed by critical social theory and holistic facilitation. IFs nominated by each of the eight settings participated in a 24-month development programme. This started with a five-day residential programme run by two EFs (BMcC & AT) followed by 20 days to work on the local implementation and evaluation of the recommendations, supported by 24 half day learning groups via teleconferencing, and 12 half days for self-directed study (38 days in total). The EFs each have over 20 years' experience of facilitation. Supplementary File 1 contains more details on the underpinning theories and activities in each intervention.

A model of co-facilitation was used in both facilitation arms where a second staff member in the organisation, a "buddy", worked with the IF, using this as a development opportunity, including taking the lead if the initial facilitator was unable to continue.

Outcomes Measures

The primary outcome was the documented percentage compliance with continence recommendations produced by the fourth International Consultation on Incontinence.[18] Percentage compliance is calculated for each resident, so is measured at the resident level.

These recommendations included 1) the resident should be actively screened for incontinence (five components), 2) a detailed assessment should be carried out (15 components), 3) an individualised treatment plan should be in place (13 components) and 4) a specialist referral should be made if needed (one component). These outcomes were assessed at baseline, then at 6, 12, 18 and 24 months. Supplementary File 2 lists all components of the continence recommendations.

Secondary outcomes included the documented incidence of level of cognitive impairment (as this influences the type of continence care the guidance recommends), depression, incontinence related dermatitis, urinary tract infections (UTIs), health related quality of life (EQ-5D[19] and IQoL[20]) and the proportion of residents in the setting with incontinence and use of pelvic floor exercises. Organisational context was assessed using the Alberta Context Tool (ACT). [21,22] The ACT data was collected from Nurses, Licenced Practical Nurses (LPN) and Health Care Assistants (HCA) at baseline in 23 of the 24 sites.

Sample Size and Power calculations: There was no information on existing compliance with the continence recommendations. We took a 50% compliance as an initial assumption. It was assumed that each setting would have 50 residents available for assessing compliance. For 90% power to detect compliance of 15% better in the intervention compared to control arm and allowing for an intra-cluster correlation of 0.01 (typically found in Primary Care Studies[23])and statistical tests carried out at the 5% level, for a cluster size of 50, seven clusters (long term care settings) were required per intervention arm. Thus 7x3 arms=21 clusters were needed. Allowing for potential attrition, this was increased to eight clusters per arm, so 24 clusters in total. This equates to 6 long term nursing care settings in each of the four countries with 50 or more residents per setting. Consent was sought at cluster and at individual level, the former before randomisation and the latter after randomisation.

Randomisation sequence generation, allocation concealment, implementation and blinding

In each country, sites were randomly allocated to one of three arms (standard dissemination, and two different intensities and kinds of a facilitation intervention), using a random sequence

generated by the statistician. A centralised randomisation point was set up by the study statistician to ensure allocation concealment. Long term care settings were enrolled by country leads for the study. The statistician was blinded to the intervention group. It was not possible to blind site staff to intervention. Research fellows who collected data from records and where necessary obtained consent from residents were blinded to the intervention group, but as discussed in the protocol[5], previous experience suggested this blinding may be inadvertently broken by the sites.

Quantitative Analysis – Statistical Methods: Data were analysed using STATA15. The primary outcome measures, percentage compliance, were analysed by fitting multi-level mixed effects linear regression models with standard errors adjusted for the clustering at the level of the nursing care setting (site level).[24] Data was collected every six months, but because the resident population was constantly changing it is necessary to consider the data as repeated cross-sectional assessments of residents in the care settings rather than longitudinal assessment of individuals within the care settings. The regression models include three independent variables: study arm (three levels), country (four levels), time period (five levels baseline, 6 months, 12 months, 18 months, 24 months), interaction terms would only be fitted if study arm main effects were significant. Intra-cluster correlation coefficients (ICC) for the baseline measurements of the compliance scores were calculated through ANOVA with adjustment for clustering and unequal cluster size. Post-estimation ICCs are calculated after fitting the regression models. Descriptive statistics, ANOVA and chi-square tests were used, where appropriate, to examine differences between groups with regard to secondary outcomes. Data were examined by an independent data monitoring committee.

Qualitative Analysis: The process evaluation data were analysed from a realist perspective[25] and are reported in the linked paper (Rycroft-Malone et al[36]).

Findings

In each country, we planned to recruit six sites (two sites per arm). This happened for Sweden and the Netherlands. However, one site in England withdrew before the study started. When no additional site was forthcoming in England within the timeframe, an additional control site was recruited in Republic of Ireland, and ethical clearance obtained. This final site had data collected up to month 18 only as there was not time to collect data at 24 months. There were thus five sites in England (with one site in the control arm) and seven in the Republic of Ireland (three sites in the control arm). Each cluster (site) received the allocated intervention and were analysed for primary and secondary outcomes.

Quantitative data were available from 2313 resident records across all time points (n=430 at baseline, n=462 at 6 months, n=497 at 12 months, n=479 at 18 months and n= 445 at 24 months after the intervention). The sample is described and then the primary outcome, compliance with the four continence recommendations is presented. The study took place between 2010 and 2013.

Description of resident sample

Most residents were included at one time point only. In all four countries, at baseline the mean age of residents varied from 82-87 years. This was almost unchanged at 24 months later (range 82-86 years). In all four countries there were more female than male residents at baseline (the percentage female in each site ranged from 60-71%), and this was similar at 24 months (range 54-80% females). At baseline the mean age of residents allocated to the three intervention groups was very similar (Control 85.34 years (s.d. 7.39); Type A 86.35 years (s.d. 7.19); Type B 83.20 years (s.d. 8.48)). The gender mix was also similar for the three intervention groups (Control 68.8% female; Type A 62.2% female; Type B 73.8%).

To understand the health status of the residents, data from the EQ-5D visual analogue scale (VAS) measure of health state that we administered at 24 months provides summary information for each intervention group (Table 1). Data at 24 months are chosen because EQ-5D-VAS was available for a higher proportion of residents than any other time point. Higher scores on a scale of 0-100 represent

better health states. Table 1 shows there was no significant difference in the mean EQ-5D scale for the intervention groups; so on average resident health status was similar in all the intervention groups. Not all residents were able to complete or have a proxy complete an EQ-5D so numbers completed are lower than total number of residents.

Table 1 here

Primary Outcomes – compliance with the four continence recommendations

(Full details of all the components of each of the four continence recommendations are available in the supplementary file 1). The ICC for percentage compliance with recommendations has been calculated from the baseline data, making allowance for both the clustering and the unequal numbers from the 24 long term care settings. The ICC for percentage compliance with recommendation 1 is 0.545 (95% CI 0.361, 0.730); for percentage compliance with recommendation 2 the ICC is 0.404 (95% CI 0.220, 0.587) and for percentage compliance with recommendation 3 ICC was 0.455 (95% CI 0.270, 0.641). These ICCs are much higher than expected and those usually found in Primary Health Care studies of 0.01[23] they are more similar to those found in some educational cluster trials.[26]

The results reported in tables 2, 4 and 6 are from fitting multi-level mixed effect linear regressions models to the compliance scores for each of the recommendations 1,2 and 3 respectively. These models account for the cluster design by treating site as a random effect and adjusting the standard error for the 24 site clusters. The model includes three independent variables: study arm (three levels – control, Type A and Type B), country (four levels – Netherlands, Sweden, Republic of Ireland and England), time period (five levels: baseline, 6 months, 12 months, 18 months, 24 months). The first level for each variable (Control arm for intervention, Netherlands for country and baseline for time) are taken as the base level and other levels are compared to this. In this model we are considering the effect of intervention allowing for country and time. The assumptions of linear

regression were examined and there was no evidence that the data failed to meet these assumptions. As a sensitivity analysis the linear regression models were also fitted omitting the country covariate and this did not change any of the findings with regard to the significance of the intervention effect.

Compliance with Recommendation 1: The resident should be actively screened for urinary incontinence

Compliance with recommendation 1 can range from 0 to 5 depending on which of five potential components of this recommendation are documented. For each component documented one point is scored, percentage compliance is a score out of 5 as a percentage. Table 2 reports the model for compliance with recommendation 1 and shows outcome scores in the intervention arms did not reach statistical significance. Country is significant with Sweden having poorer compliance (a negative coefficient) compared to the Netherlands. Ireland and England had significantly better compliance than the Netherlands (positive coefficients). The 12 and 24 month data collection parameters were significant, but the other points were not significantly different to baseline. The post-estimation ICC following the fitting of this model for compliance with recommendation 1 is 0.091. Table 3 shows the mean percentage for each intervention group at each time point, showing the small increase in percentage compliance score for type A and type B intervention up to 12 months, though as the regression model indicates there is no significant difference in the study arms over the duration of the study.

Table 2 here

Table 3 here

Compliance with Recommendation 2: A detailed assessment should be carried out

There are 15 items in the detailed assessment, so scores can range from 0-15 for recommendation 2. Percentage compliance is a score out of 15 as a percentage.

Table 4 reports the fitted model for compliance with recommendation 2. The intervention is not effective; neither the Type A facilitation or Type B facilitation interventions had significant coefficients. Ireland was significantly different having higher compliance with recommendation 2, but the coefficients for the other countries were not significant, so England and Sweden are not significantly different to the Netherlands after allowing for time point and intervention group. The 24 month data collection parameter is significant, with increased compliance by 24 months, but the other points are not significantly different to baseline. The post-estimation ICC following the fitting of this model for compliance with recommendation 2 is 0.351. Table 5 shows mean percentage compliance score for recommendation 2 by intervention group. Mean percentage compliance was low at baseline, in all groups, but improved by 24 months in the Type A and Type B intervention groups.

Table 4 here

Table 5 here

Compliance with Recommendation 3: An individualised treatment plan should be in place

A score from 0 to 13 is possible for compliance with recommendation 3. Percentage compliance is a score out of 13 as a percentage.

Table 6 reports the fitted model for compliance with recommendation 3. The intervention was not effective, neither the Type A facilitation or Type B facilitation interventions had significant coefficients. All country parameters were significant with Sweden, Ireland and England all having significantly higher compliance with recommendation 3 than the Netherlands. All time points were significant, and the parameter value increased for each successive time period, thus suggesting improvement over time in compliance with recommendation 3. This suggests learning over time in all countries but no significant difference in the effectiveness of the three study interventions. The post-estimation ICC following the fitting of this model for compliance with recommendation 3 is

0.126. Table 7 shows mean percentage compliance for recommendation 3 by intervention group. It can be seen that all three groups appear to improve over time, with little difference between the interventions as indicated by the regression model.

Table 6 here

Table 7 here

Recommendation 4: Specialist referral should be made if necessary

There were very few specialist referrals made and in the data collection it was not always clear whether a lack of documentation meant no referral was made or whether a referral was not necessary. It is therefore difficult to fully assess compliance with this guideline. However, the level of referral was so low that it is very unlikely that study arm has a significant impact on compliance with this recommendation. In only 4% of residents was a referral recommended. Although these referrals were recorded as specialist referrals, 17 were to a general practitioner (family doctor) and 6 to an unknown specialist. There were only 11 referrals to a continence specialist nurse and six referrals to urology.

In summary, for the primary outcome (documented compliance score or percentage compliance with continence recommendations) there was no significant difference between study arms; all study arms improved over time in all countries.

Secondary (clinical) outcomes

These data are being considered as two cross-sectional reviews of the resident populations in the long term care settings as there are very few individual residents included at both baseline and 24 months data collection. At 24 months there was no significant difference between the three intervention groups with regard to the proportion of residents who had no documented record of the assessment of cognition (p=0.076 from chi-square test). At 24 months there was a significant difference between the three intervention groups with regard to the proportion of residents who

had no documented record of the level of cognitive impairment (p<0.001 from chi-squared test), the proportion being higher in the control group than in the Type A and Type B groups. At 24 months there was a significant difference between the three intervention groups with regard to the proportion of residents who had no documented record of the assessment of depression (p=0.017 from chi-squared test), the proportion being higher in the control group than in the Type A and Type B groups. At 24 months there was no significant difference between the three intervention groups with regard to the proportion of residents who had no documented record of the assessment of incontinence related dermatitis (p=0.479 from chi-square test).

Between baseline and 24 months there was a statistically significant decrease in the proportion of residents who had no documented record of an assessment of cognition in Type B facilitation (p<0.001) but no significant change for Type A; there was a significant decrease in the proportion of residents who have no documented record of the level of cognitive impairment in intervention Type A (p<0.001) and Type B (p<0.001); **t**here was a significant reduction in the percentage of residents who had no documentation of assessment of depression in the Type A (p<0.001) and Type B (p<0.001) groups. There was a significant decrease in the percentage of residents who had no documentation of incontinence associated dermatitis between baseline and 24 months in the Type A (p<0.001) and Type B (p<001) groups. There was no significant improvement in the control group for any of the secondary outcomes.

Whether the impact of urinary incontinence on quality of life been assessed was not documented for the majority of residents. It was not assessed more than seven times in any group, so this was not explored further. Very few UTIs were documented. In the month prior to the baseline data collection only 15 UTIs were recorded in all countries, decreasing to only seven at the 24 month data collection point. No further analysis was done.

It was not possible to reliably calculate the proportion of residents in each long term care setting with incontinence, thus no further analysis was done. At baseline pelvic floor exercises were not

used with any residents and at 24 month follow up pelvic floor exercises were only used with 3 residents. With such low numbers, no further exploration of this is sensible.

In summary, for secondary outcomes, both the facilitation intervention groups (Type A and Type B) showed significantly better documentation of three outcomes: the level of cognitive impairment, depression and incontinence associated dermatitis between baseline and 24 months, and this improvement did not occur in the standard dissemination (control) group. Clinically this change was not large, and a substantial proportion of residents still had no documented assessment of level of cognitive impairment (68% in Type A and 65% in Type B) depression (61% in Type A and 65% in Type B) and incontinence associated dermatitis (66% in Type A and 73% in Type B).

There was a large amount of missing data on the Urinary Incontinence Quality of Life (I-QoL) outcome measure[20] as residents found it too much to complete, so this is not reported further. It had been planned to report length of stay data, but it was not possible to collect this data consistently across all sites, so it is not reported further.

Health economics

Health Economic analysis was undertaken, but since there was no significant difference in the primary outcome between the intervention groups, these data are not presented here in detail because the cost analysis showed that, as expected, standard dissemination would be the least costly intervention to implement. (see supplementary file 3 for intervention cost tables)

Alberta Context Tool

For all concepts, higher scores represent a better work context. All responses for a site (Nurse, LPN, HCA) were considered together to provide an overall picture of the site. The questionnaire completed by Nurses, LPN's and HCAs were identical except with regard to informal interactions in which the HCA group had one less question (9) than the other groups of staff who had 10 questions in this section.

Table 8 here

Table 8 shows for each concept the mean score given by all staff rating a site within the intervention arm. Formal interactions are notably lower than other scores. The largest differences are for structural and electronic resources and for organisational slack - space. On the basis of the similarity of these mean scores we conclude the study groups were similar with regard to ACT concepts.

Discussion

The 12 months Type A and the 24 months Type B facilitation interventions did not have different levels of impact on documented compliance with recommendations. It was thus not possible to identify the type and "dose" of facilitation that worked best within the highly varied contextual conditions identified in this study. In addition, the process evaluation revealed important issues about the models of facilitation used and the characteristics of the facilitators (see linked paper Harvey et al[37]).

So why was it that the facilitation intervention did not make a statistically significant difference to the documented implementation of continence recommendations? Was an element of the PARIHS framework, facilitation, purported to be necessary for getting research evidence used in practice, actually not so important? Other research has found some type of help with getting research implemented does make a difference.[27,28] Baskerville et al's[29] systematic review of practice facilitation in primary care suggests facilitation improves uptake of clinical practice guidelines by nearly three times. A facilitation intervention was found to reduce neonatal mortality by 50%.[30] Although the facilitation not working in this study is a possible explanation and the high ICCs meant the study was underpowered, the process evaluation qualitative research evidence (linked paper Rycroft-Malone et al[36]) suggested this was not the most likely explanation. It may be facilitation works differently along the continuum of context. It could be that using only documented evidence of compliance with the recommendations under- estimated what might have happened in practice

but was not documented. A lack of intervention fidelity is another possible explanation, and this is addressed in the process evaluation paper (linked paper Rycroft-Malone et al[36]).

Although the intervention groups improved, it was not possible to say the improvement was due to the intervention as the control group also improved. We do not know why this was, but it could be that for control sites, being in the study, including 6 monthly follow-ups for two years, was enough of an incentive to improve. However, the qualitative data suggests for most control sites they did not use the written recommendations or the implementation guide. One site mentioned to the researcher that they checked their documentation and practice knowing the researcher would be visiting, and thus even collecting follow-up data in the control group can be seen as having an effect.

Etheridge et al[31] concluded that four active ingredients were required to effect change in long term care settings: urgency, solidarity, intensity and accumulation. The continence programme they reviewed failed and one of the reasons they identified may also apply in our study: there was no buy-in from participants. Although all sites agreed to take part in the study, the topic area and the intervention were already decided. In addition, participants changed during the study, so, for example, as managers changed, new managers did not necessarily see this study as a priority, thus reducing even further the extent of organisational buy-in and support (see linked paper[36]).

The proposition that underpins the PARIHS framework is that successful implementation is a function of the nature of the evidence being implemented, the context into which it is being implemented and appropriate facilitation to help people implement the evidence. There was no weighting given to these three aspects of evidence, context and facilitation. This research suggested that facilitation with one or two people in a team may not easily overcome contextual factors. The level of experience and expertise of the IF, and relationship of the IF to managers in the setting may be more important[32] as may unravelling how facilitation and context interact.

It was not possible to identify a “good enough” model of facilitation that affected the primary outcome (documented compliance with continence recommendations) and could address the

different contexts. Facilitation did however result in some identifiable practice changes (e.g. new assessment processes, new forms, and awareness of the impact of incontinence on residents).

It may be that in practice, tailoring the type of facilitation to both the setting and the internal facilitator is important. Just how one could map the contextual characteristics to a type of facilitation and to type of internal facilitator would need further evaluation. Van der Zijpp et al,[32] part of this study, argued the interactions between managerial leaders and IFs were important, summarised by three themes: realising commitment, negotiating conditions and encouraging to keep the momentum going. The reciprocal relationships between managers and IFs influenced the process of implementation and future interventions should target managers in a focused way. In studies that evaluate implementation of complex interventions such as facilitation, it may be appropriate to adopt a theoretical perspective on fidelity, focusing on the intended mechanisms of the intervention. For example, in this study the theory of Type A facilitation required IFs to develop skills and confidence in audit and feedback. Achieving this mechanism, even if it meant IFs needed varying levels of external facilitation, would demonstrate theoretical fidelity. This type of approach has been proposed in public health[33] and is discussed in more depth in the linked papers (Rycroft-Malone et al[36], Harvey et al[37]).

ACT considers organisational concepts as a unit-based score. In this study these were considered as site level variables. Mean baseline and follow-up mean scores were compared with either an ANOVA where multiple time points were available or with a t-test when only one follow-up time point was available. There were very few changes that were significant. We are thus not confident to make any claims about the effects of the intervention on organisational culture as assessed with ACT. Possible explanations for this include the organisations were stable and at site level the concepts were unaffected by the interventions.

Limitations

In reality, the planned interventions did not always work as originally envisaged, as revealed by the process evaluation[36] and our analysis of the facilitation intervention[37]. This was for several reasons, relating to: initial selection and preparation of the IFs; engagement in the facilitation intervention; ability to progress according to plan. The linked papers illustrate the issues that compromised the fidelity of the intervention (Rycroft-Malone et al[36] linked paper). It was also challenging to recruit resident participants in some homes, so we had fewer than planned. In addition, although each of the long term settings had agreed to take part, for individual staff within the home it was not necessarily a priority. The unexpectedly high ICC meant the study was underpowered. Although we felt the ICC we used in the sample size calculation was reasonable, in planning future cluster RCTs with a more educational focus, it is important to be aware that not all ICCs will be as low as those reported for recent primary care trials. [23] In the design of the study it was assumed that there would not be large country differences regarding compliance with the recommendations, hence it would be viable to have a small number of sites from each country in each study arm. In practice it appears the countries are behaving differently, but the study was not powered to investigate within country effect of the different interventions on the primary outcome.

Conclusions

Pressman & Wildavsky[34] a long time ago reported that "the study of implementation requires understanding that apparently simple sequences of events depend on complex chains of reciprocal interaction" (pxvii) and referred to the complexities of implementation as "the lumpy stuff of life"[35] (p165). This study supports those assertions.

This was the first cross European randomised controlled trial with embedded process and economic evaluations that sought to test different methods of facilitation. There was no significant difference in the primary outcome between any of the three study arms. It found both models of facilitation were broadly viable but were not significantly better than a control in improving documented compliance with recommendations to promote continence. Contextual issues were not always

overcome by the approaches to facilitation adopted in this study as our linked papers demonstrate (Rycroft-Malone et al, Harvey et al[36,37]).

Declarations

Ethics approval and consent to participate

Ethical Committee approval was obtained in England (10/WSE04/20), Sweden (2009/1806-31/2), and Republic of Ireland (ECM4(u)02/03/10). In the Netherlands, the researchers followed advice to get permission from either an ethical committee at site level, or where this did not exist, from a scientific or residents committee at the site (HAZ-11087777-JGS). Research Governance approval was also obtained in England and permission to collect data at the sites obtained in Sweden and Republic of Ireland.

Consent for publication

Consent form allowed the use of anonymised quotations in publications.

Availability of data and materials

The datasets generated or analysed during the current study are not publicly available because consent to make data publicly available was not part of the consent by participants.

Competing interests – We acknowledge that CE is involved in the development of the Alberta Context Tool and AK GH JRM BMcC KS and AT have all been involved in the development of the PARIHS framework.

Funding

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 223646.

Authors' contributions

All authors read and approved the final manuscript.

KS (principal investigator) led the application for funding. She contributed to the overall design of the study, designed the cluster RCT aspect of the study, and contributed to the analysis. She led the writing of this paper.

JRM (collaborator) participated in designing the study. She led the design of the evaluation package and was country co-ordinator for England. She wrote the process evaluation aspects of this paper and reviewed the manuscript critically for important intellectual content.

KC (collaborator) participated in designing the study and reviewed the manuscript critically for important intellectual content. She was country co-ordinator for the Netherlands.

NC (statistician) advised on study design, contributed to the analysis plan and undertook the analysis within the cluster RCT, and reviewed the manuscript critically for important intellectual content.

RTE and CJ (health economists) were responsible for the economic evaluation study design and reviewed the manuscript critically for important intellectual content.

ACE (research fellow) participated in the design and analysis of the evaluation package, collected data in Sweden, contributed to the analysis and reviewed the manuscript critically for important intellectual content.

CAE (collaborator) participated in study design and coordinated the use of the Alberta context tool including its translation into Swedish and Dutch. She reviewed the manuscript critically for important intellectual content.

CH (research fellow) participated the design of the process evaluation and associated data collection tools, the development of the economic evaluation, and was responsible for the day to day running

of the process evaluation, contributed to the data collection in England and the analysis and reviewed the manuscript critically for important intellectual content.

GH (collaborator) participated in the design of the overall study and in the design of the facilitator intervention in particular. She co-led Type A facilitation work package and reviewed the manuscript critically for important intellectual content.

AK (collaborator) participated in the design of the overall study and in the design of the facilitator intervention in particular. She co-led Type A facilitation work package and reviewed the manuscript critically for important intellectual content.

BMcC (collaborator) participated in the design of the overall study and in the design of the facilitation intervention in particular. He co-led Type B facilitation work package and reviewed the manuscript critically for important intellectual content. He was also Country Coordinator for Ireland.

CM (research fellow) participated in the design of the RCT and associated data collection tools, participated the collection of data in England, contributed to the analysis and reviewed the manuscript critically for important intellectual content.

AT (collaborator) participated in the design of the overall study, especially the Type B facilitation intervention. She co-led Type B facilitation work package and reviewed the manuscript critically for important intellectual content.

PS and CMcC (research fellows) participated the collection of data in Republic of Ireland, contributed to the analysis and reviewed the manuscript critically for important intellectual content.

TN and TvdZ (research fellows) participated the collection of data in Netherlands, contributed to the analysis and reviewed the manuscript critically for important intellectual content.

LW (collaborator) participated in the design of the overall study and in the design of the intervention evaluation in particular. He reviewed the manuscript critically for important intellectual content and was Country Coordinator for Sweden.

Acknowledgements

Thank-you to all those who were involved in this study as participants or who advised us on the study.

Note to editorial assistant:

Please note that references to Rycroft Malone et al and Harvey et al ***as linked papers*** *refer to two papers already accepted by Implementation Science which will be published together with this paper. I will reference those fully when I have this information from Implementation Science. Reference 36 and 37 in the list.*

References

1. Grol R. Success and failures in the implementation of evidence-based guidelines for clinical practice. Medical Care. 2001; 39(8 Suppl 2):1146–1154

2. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of care delivered to adults in the United States. New England Journal of Medicine. 2003; 348(26):2635–2645.

3. Katikreddi SV, Higgins M, Bond L, Bonnell C, Macintyre S. How evidence based is English public health policy? BMJ 2011;343:d7310

4. Kitson A, Rycroft-Malone J, Harvey G, McCormack B, Seers K, Titchen A. Evaluating the successful implementation of evidence into practice using the PARIHS framework: Theoretical and practical challenges, Implementation Science, 2008; 3:1 doi:10.1186/1748-5908-3-1

5. Seers K, Cox K, Crichton NJ, Tudor-Edwards R, Eldh A, Estabrooks CA, Harvey G, Hawkes C, Kitson A, Linck P, McCarthy G, McCormack B, Mockford C, Rycroft-Malone J, Titchen A, & Wallin, L. FIRE (Facilitating Implementation of Research Evidence): a study protocol. Implementation Science 2012; 7: 25 doi:10.1186/1748-5908-7-25

6. Seers K, Cox K, Crichton NJ, Tudor-Edwards R, Eldh A, Estabrooks CA, Harvey G, Hawkes C, Kitson A, Linck P, McCarthy G, McCormack B, Mockford C, Rycroft-Malone J, Titchen A, & Wallin L. (2013) Final Report Summary - FIRE (Facilitating Implementation of Research Evidence) http://cordis.europa.eu/result/rcn/149765_en.html. Accessed 03/04/18

7. Eccles MP, Armstrong D, Baker R, Cleary K, Davies H, Davies S, Glasziou P, Ilott I, Kinmonth AL, Leng G, Logan S, Michie S, Rogers H, Rycroft-Malone J, Sibbald B. An implementation research agenda. Implementation Science 2009; 4:18 doi:10.1186/1748-5908-4-18

8. Grimshaw JM, Thomas RE, MAcLennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles M, Matowe L, Shirran L, Wensing M, Dijkstra R, Donaldson C. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technology Assessment 2004; 8:6

9. Thompson DS, Estabrooks CA, Scott-Findlay S, Moore K, Wallin L. Interventions aimed at increasing research use in nursing: a systematic review. Implementation Science. 2007; 2:15. http://www.implementationscience.com/content/pdf/1748-5908-2-15.pdf

10. Medves J, Godfrey C, Turner C, Paterson M, Harrison M, MacKenzie L, Durando P. Systematic review of practice guidelines dissemination and implementation strategies for healthcare teams and team-based practice. Int J Evidence Based Healthcare. 2010; 8:78-89 doi:10.1136/bmj.a2390

11. Dogherty EJ, Harrison MB, Graham ID. Facilitation as a role and process in achieving evidence-based practice in nursing: a focused review of concept and meaning. Worldviews Evidence Based Nursing. 2010; 7(2):76–89.

12. Berta W, Cranley L, Dearing JW, Dogherty EJ, Squires JE and Estabrooks CA. Why (we think) facilitation works: insights from organizational learning theory. Implementation Science. 2015; 10:141 DOI 10.1186/s13012-015-0323-0

13. Brittain KR, Shaw C. The social consequences of living with and dealing with incontinence – a carers perspective. Social Science and Medicine. 2007; 65:1274-1283

14. Ouslander JG & Schnelle JF. Incontinence in the nursing home. Ann Intern Med. 1995;122(6):438-449

15. National Guideline Clearinghouse. Guideline for Urinary Incontinence in the long term setting. American Medical Directors Association. Agency For Healthcare Research and Quality. 2012; http://www.guideline.gov/content.aspx?id=45526. Archived. https://www.guideline.gov/summaries/archive/withdrawn/771#jumptoU. Accessed 3/4/18

16. Perry L, Bellchambers H, Howie A, Moxey A, Parkinson L, Capra S, Byles J. Examination of the utility of the Promoting Action on Research Implementation in health Services framework for implementation of evidence based practice in residential aged care settings. Journal of Advanced Nursing. 2011; 67(10):2139-2150 ;67(10):2139-50

17. Rycroft-Malone J, Seers K, Crichton N, Chandler J, Hawkes C, Allen C, Bullock I, Strunin L. A pragmatic randomised controlled trial evaluating three implementation interventions. Implementation Science. 2012; 7:80.

18. DuBeau CE, Kuchel GA, Johnson T, Palmer MH, Wagg A. Committee 11. Incontinence in the frail elderly. In: Abrams P, Cardozo L, Khoury S & Wein A (eds) . Incontinence. 4th International Consultation on Incontinence. Health Publications Limited. 2009; pp961-1024 and pp1796-1789.

http://www.icsoffice.org/Publications/ICI_4/book.pdf;

https://pdfs.semanticscholar.org/0b60/32e462e6b0fd6f9a255d0035abe461f77faf.pdf

Accessed 3/04/18

19. Devlin N, Parkin D, Browne J. (2010). Using the EQ-5D as a performance measurement tool in the NHS. *Health Economics* 19(8):886-905.

20. Patrick DL, Martin ML, Bushnell DM, Yalcin I, Wagner TH, Buesching DP. Quality of life with women with urinary incontinence: further development of the incontinence quality of life instrument (I-QOL). Urology. 1999; 53:71-76

21. Estabrooks CA,Squires JE, Cummings GG, Birdsell JM, Norton PG. Development and assessment of the Alberta Context Tool. BMC Health Services Research. 2009; **9**:234 **DOI:** 10.1186/1472-6963-9-234

22. Squires JE, Hayduk L, Hutchinson AM, Mallick R, Norton PG, Cummings GG, Estabrooks CA. Reliability and validity of the Alberta Context Tool (ACT) with professional nurses: findings from a multi-study analysis. PLoS One. 2015; 10(6):e0127405

23. Campbell MK, Mollison J, Steen N, Grimshaw JM &Eccles M (2000) Analysis of cluster randomized trials in primary care: a practical approach. Family Practice 17:192-196.

24. Donner A & Klar N (2000) Design and analysis of cluster randomized trials in health research. Arnold. London.

25. Pawson R. The Science of Evaluation: A Realist Manifesto. Sage, London. 2013

26. Brooks G, Burton M, Cole P, Miles J, Torgerson C & Torgerson D (2008) Randomised controlled trial of incentives to improve attendance at adult literacy classes. Oxford Review of Education 34:493-504.

27. Eriksson L, Huy TQ, Duc DM, Ekholm Selling K, Hoa DP, Thuy NT, Nga NT, Wallin L. Process evaluation of a knowledge translation intervention using facilitation of local stakeholder groups to improve neonatal survival in the Quang Ninh province, Vietnam. Trials. 2016;17:23.

28. McCormack B, Rycroft-Malone J, DeCorby K, Hutchinson AM, Bucknall T, Kent B, Schultz A, Snelgrove-Clarke E, Stetler C, Titler M ,Wallin L, Wilson L. A realist review of interventions and strategies to promote evidence-informed healthcare: a focus on change agency. Implementation Science. 2013; **8**:107 DOI**:** 10.1186/1748-5908-8-107

29. Baskerville B, Liddy C, Hogg W. Systematic Review and Meta-Analysis of Practice Facilitation within Primary Care Settings. Ann Fam Med. 2012;37(8):63-74

30. Persson LÅ, Nguyen TN, Målqvist M, Dinh Thi Phuong H, Eriksson L, Wallin L, Tran QH, Duong MD, Tran VT, Vu Thi Thu T, Ewald U. Effect of facilitation of local maternal-and-child stakeholder groups on neonatal mortality. The NeoKIP cluster-randomised trial in Quang Ninh province, Vietnam. PLOS Medicine 2013;10(5). http://dx.doi.org/10.1371/journal.pmed.1001445

31. Etheridge F, Couturier Y, Denis JL,Tremblay L, & Tannenbaum C. Explaining the success or failure of quality improvement initiatives in long term care organisations from a dynamic perspective. Journal of Applied Gerontology. 2014; 33(6):672-689

32. Van der Zijpp TJ, Niessen T, Eldh A, Hawkes C, McMullan C, Mockford C, Wallin L, McCormack B, Rycroft-Malone J & Seers K. A bridge over turbulent waters - illustrating the interaction between managerial leaders and facilitators when implementing research evidence. Worldviews on Evidence-Based Nursing. 2016;13(1):25-31. doi: 10.1111/wvn.12138

33. Hawe P. Lessons from complex interventions to improve health. Annual Review of Public Health. 2015; **36:**307-323

34. Pressman JL & Wildavsky A. Implementation. How great expectations in Washington are dashed in Oakland or Why it's amazing that federal programs work at all this Being a Saga of the Economic Development Administration as Told by Two Sympathetic Observers Who Seek to Build Morals on a Foundation of Ruined Hopes. University of California Press, Berkeley. 1973

35. Pressman JL & Wildavsky A. Implementation. How great expectations in Washington are dashed in Oakland; or Why it's amazing that federal programs work at all this Being a Saga of the Economic

Development Administration as Told by Two Sympathetic Observers Who Seek to Build Morals on a Foundation of Ruined Hopes. 3rd edition. University of California Press, Berkeley. 1984

36. Rycroft-Malone et al (linked paper) Jo Rycroft-Malone; Kate Seers; Ann Catrine Eldh; Karen Cox; Nicola Crichton; Gill Harvey; Claire Hawkes; Alison Kitson; Brendan George McCormack; Christel McMullan; Carole Mockford; Theo Niessen; Paul Slater; Angie Titchen; Teatske van der Zijpp; Lars Wallin. A realist process evaluation within the Facilitating Implementation of Research Evidence (FIRE) cluster randomised controlled international trial: an exemplar. Accepted by Implementation Science.

37. Harvey et al (linked paper) Harvey G McCormack B Kitson A Lynch E Titchen A. Designing and implementing two facilitation interventions within the 'Facilitating Implementation of Research Evidence (FIRE)' study: A qualitative analysis from an external facilitators' perspective. Accepted by Implementation Science

List of Abbreviations

EF – External Facilitator

FIRE - Facilitating Implementation of Research Evidence

IF – Internal Facilitator

PARIHS - Promoting Action on Research Implementation in Health Services

UI – Urinary incontinence

UTI – Urinary Tract Infection

Table 1: Summary statistics for EQ-5D-VAS scale for each intervention group

Intervention	Number of residents with completed scale	Mean (SE robust)*	95% CI for mean	Range
Standard Dissemination (Control)	109	54.2 (4.737)	44.35, 64.00	0, 100
Type A	113	59.2 (4.325)	50.19, 68.13	0, 100
Type B	124	55.6 (2.918)	49.57, 61.67	0, 90

**SE robust allows for the clustering, and ANOVA allowing for clustering to compare the three means, gave p=0.34*

Table 2: Multilevel mixed effect linear regression model – Percentage compliance with recommendation 1 (The resident should be actively screened for urinary incontinence), with adjustment of standard errors to allow for clustering.

	Coefficient	Std. Err.	z	P -value	95% confidence interval
Type A	2.9293	3.1298	0.94	0.349	-3.2049, 9.0635
Type B	-4.1688	3.6966	-1.13	0.259	-11.4141, 3.0765
Sweden	-31.0840	4.0940	-7.59	0.000	-39.1082, -23.0599
Ireland	13.8449	4.5226	3.06	0.002	4.9808, 22.7091
England	10.3152	5.0742	2.03	0.042	0.3699, 20.2604
+ 6 months	0.1104	2.8436	0.04	0.969	-5.4630, 5.6837
+ 12 months	12.9885	4.4264	2.93	0.003	4.3130, 21.6641
+18 months	4.9052	3.3204	1.48	0.140	-1.6025, 11.4130
+24 months	9.3776	4.3632	2.15	0.032	0.8259, 17.9292
Constant	33.7259	4.2278	7.98	0.000	25.4396, 42.0122

N=2313; Model fit: Wald $\chi^2(9)=1970.23$, $p<0.001$; Post-estimation ICC 0.0910 (se 0.0219)

Table 3: Mean percentage compliance with recommendation 1 by intervention group for each time point

Intervention Group	Mean score				
	Baseline	**6 month**	**12 month**	**18 month**	**24 month**
Control	28.4	22.3	29.2	27.0	23.2
Type A	19.2	21.5	38.8	30.6	35.5
Type B	14.1	17.0	44.4	23.4	28.7

N=2313 residents are included in this analysis

Table 4: Multilevel mixed effect linear regression model - Percentage compliance with recommendation 2 (A detailed assessment should be carried out), with adjustment of standard errors to allow for clustering.

	Coefficient	**Std. Err.**	**z**	**P -value**	**95% confidence interval**
Type A	5.6514	4.0014	1.41	0.158	-2.1912, 13.4941
Type B	3.7903	4.4807	0.85	0.398	-4.9917, 12.5724
Sweden	-1.9108	2.7374	-0.70	0.485	-7.2760, 3.4545
Ireland	14.9312	3.6627	4.08	0.000	7.7524, 22.1099
England	11.7997	7.0278	1.68	0.093	-1.9745, 25.5738
+ 6 months	-0.2220	1.2763	-0.17	0.862	-2.7235, 2.2794
+ 12 months	3.3623	2.1118	1.59	0.111	-0.7767, 7.5014
+18 months	-0.0031	1.6463	-0.00	0.998	-3.2298, 3.2235
+24 months	4.4827	2.1665	2.07	0.039	0.2364, 8.7290
Constant	30.1617	3.3204	9.08	0.000	23.6538, 36.6696

N=2313; Model fit: Wald $\chi^2(9)=64.76$, $p<0.001$; Post-estimation ICC 0.3517 (se 0.0758)

Table 5: Mean percentage compliance with recommendation 2 by intervention group for each time point

Intervention Group	Mean score				
	Baseline	**6 month**	**12 month**	**18 month**	**24 month**
Control	37.5	34.6	36.5	34.1	34.4
Type A	34.6	35.1	45.1	39.7	44.6
Type B	35.3	34.8	43.2	38.2	45.9

N=2313 residents are included in this analysis

Table 6: Multilevel mixed effect linear regression model – Percentage compliance with recommendation 3 (An individualised treatment plan should be in place), with adjustment of standard errors to allow for clustering.

	Coefficient	**Std. Err.**	**z**	**P -value**	**95% confidence interval**
Type A	0.3391	4.0168	0.08	0.933	-7.5336, 8.2118
Type B	1.0372	3.0579	0.34	0.734	-4.9562, 7.0305
Sweden	23.7959	1.9736	12.06	0.000	19.9278, 27.6640
Ireland	24.5448	3.9162	6.27	0.000	16.8692, 32.2204
England	15.3118	3.8489	3.98	0.000	7.7681, 22.8555
+ 6 months	9.8431	4.1862	2.35	0.019	1.6382, 18.0479
+ 12 months	14.2761	3.7488	3.81	0.000	6.9285, 21.6237
+18 months	15.9399	3.7804	4.22	0.000	8.5305, 23.3494
+24 months	19.9791	3.3984	5.88	0.000	13.3183, 26.6399
Constant	6.5831	3.0927	2.13	0.033	0.5216, 12.6446

N=2313; Model fit: Wald $\chi^2(9)=387.72$, $p<0.001$; Post-estimation ICC 0.1265 (se 0.0502)

Table 7*:* Mean percentage compliance with recommendation 3 by intervention group for each time point

Intervention Group	Mean score				
	Baseline	**6 month**	**12 month**	**18 month**	**24 month**
Control	20.9	30.8	40.9	45.0	48.9
Type A	23.8	32.2	41.9	42.7	45.2
Type B	26.7	38.6	40.7	41.1	45.9

N=2313 residents are included in this analysis

Table 8: Mean scores on ACT concepts by intervention group at baseline (N=725 staff are included in this analysis)

ACT Concept*	**Number of items**	**Range for score**	**Control sites Mean (SD)**	**Type A sites Mean (SD)**	**Type B sites Mean (SD)**
Leadership#	6	1-5	3.6 (0.81)	3.7 (0.82)	3.7 (0.76)
Culture#	6	1-5	3.9 (0.65)	3.9 (0.57)	3.9 (0.61)
Feedback#	6	1-5	3.5 (0.79)	3.4 (0.82)	3.4 (0.85)
Formal Interactions~	4	0-4	1.3 (1.14)	1.1 (1.08)	1.2 (1.13)
Informal Interactions~	9 or 10	0-10	3.5 (2.11)	3.2 (2.08)	3.3 (2.04)
Connections (Social Capital)#	6	1-5	4.0 (0.67)	3.8 (0.59)	3.9 (0.59)
Structural & electronic resources~	11	0-11	3.1 (2.34)	3.4 (2.14)	2.8 (1.89)
Organisational Slack-Staffing#	3	1-5	2.7 (1.13)	2.8 (1.09)	2.6 (1.00)
Organisational Slack-Space#	3	1-5	3.6 (1.01)	3.1 (1.14)	3.3 (1.10)
Organisational Slack-Time#	4	1-5	2.8 (0.69)	2.8 (0.70)	2.8 (0.74)

*Definitions of ACT concepts and scaling are provided[21, 22] and relevant papers are listed at https://trecresearch.ca/alberta_context_tool. # scaled; ~count based

17th October 2018

Dear Editor

We have addressed all the comments from the reviewers and are pleased to submit this revised manuscript for your consideration.

Kind Regards

Kate Seers