The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-003713
Article Type:	Research
Date Submitted by the Author:	02-Aug-2013
Complete List of Authors:	Douet, Lisa; University of Southampton, National Institute for Health Research Milne, Ruairidh; University of Southampton, Wessex Institute Anstee, Sydney; University of Southampton, National Institute for Health Research Habens, Fay; University of Southampton, National Institute for Health Research Young, Amanda; University of Southampton, NETSCC Wright, David; University of Southampton, National Institute for Health Research
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Research methods
Keywords:	Clinical trials < THERAPEUTICS, AUDIT, STATISTICS & RESEARCH METHODS

The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Dr Lisa Douet[a]
Senior Research Fellow

Professor Ruairidh Milne[b]
Professorial Fellow in Public Health

Sydney Anstee[a]
Research Fellow

Fay Habens[a]
Programme Manager

Amanda Young[a]
Senior Research Fellow

Dr David Wright[a]
Senior Research Fellow

[a]National Institute for Health Research
Evaluation, Trials and Studies Coordinating Centre (NETSCC)
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

[b]Wessex Institute
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

Correspondence to: l.douet@southampton.ac.uk

Acknowledgements:
We would like to acknowledge Professor James Raftery and the Metadata team for providing the database and the trial details used in the study and for Professor Paul Glasziou for his advice during the study.

Ethic statement:
This study did not require ethics approval as no data patient or clinical data was required.

Funding statement:
This study was supported by the NIHR Evaluation, Trials and Studies Coordinating Centre through its Research on Research Programme. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Department of Health, or of NETSCC.

Contribution of Authors:
The study was conceived and designed by LD, RM, SA, and FH, and undertaken by LD, SA and FH; AY and DW supported the data analysis. All authors read and approved the final manuscript. LD is guarantor of the study

Data sharing: Data on the included trials are available on request from the corresponding author.

Competing Interests: The authors have no competing financial interests; however, all of the authors are employed by the University of Southampton to work at least part time for NETSCC. In particular: RM is employed as the Head of NETSCC and has worked for NETSCC (and its predecessor organisation) in senior roles on and off since 1996. He was an editor of the Health Technology Assessment journal (1997-2007) and a founder editor for other journals in the new NIHR Journals Library (2011-12).

Abstract word count: 239
Main text word count: 2480

Abstract:

Objectives

The objective of this study is to assess whether NIHR HTA funded randomised controlled trials (RCTs) published in *Health Technology Assessment* Journal were described in sufficient detail to replicate in practice.

Methods

A published checklist for assessing intervention descriptions was applied to NIHR HTA funded RCTs published in *Health Technology Assessment*. The checklist was piloted twice on a sample of 10 reports. Kappa scores were generated to assess agreement in the checklist application. The checklist was modified and applied to all 98 NIHR HTA funded single trial RCTs published in the Journal up to March 2011. Three assessors independently applied the checklist. Disagreements in scoring were discussed in the team; differences were then explored and resolved.

Results

Components of the intervention description were missing in 68 / 98 (69.4%) reports. Baseline characteristics and descriptions of settings had the highest levels of completeness with over 90% of reports complete. Reports were less complete on patient information with 58.2% of the journals having an adequate description. Intervention descriptions were more not significantly more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively. Only 27.3% of RCTs with psychological interventions were deemed to be complete, although numbers were too small for differences to be significant statistically.

Conclusions

Ensuring the replicability of study interventions is an essential part of adding value in research. Research funders need to ensure transparency and completeness in the reporting of interventions.

Article summary

Article focus

- It has previously been suggested that over 50% of intervention descriptions are not sufficiently described. This article investigates the description of interventions within the NIHR HTA journal series.

Key messages

- Only 30.6% (30/98) of studies with a single trial published in the NIHR HTA Journal have a full description of the intervention
- The unlimited word count of the NIHR journal series does not affect completeness of an intervention description

Strengths and limitations

- An externally produced checklist was applied to all RCTs publishing in the NIHR HTA monograph
- The small sample size for a number of assessments is very small

Introduction:

A recent publication by Glasziou and Chalmers has suggested that as much as 85% of the US$100 billion spent on health research worldwide each year is potentially wasted due to four key problems of knowledge production and dissemination. These four areas include: 1) ensuring the right research questions are asked; 2) ensuring that study designs are appropriate and are of methodological quality; 3) ensuring the findings from funded research are available in the public domain; 4) ensuring that funded research is unbiased and usable[1].

Several studies have specifically assessed the fourth area of waste by exploring the quality and usability of publications from funded health research. This is a key concern considering the role effective summaries of evidence have in facilitating knowledge transfer and enhancing the uptake of findings in clinical practice. Whilst it is recognised that trial registration databases and scientific journals are restrictive in terms of word allowance, various strategies have been proposed to improve the reporting of interventions in published trials, including an 'intervention bank' to include manuals and fidelity tools linked to trial registration numbers'[2].

Studies have highlighted concerns about the descriptions of interventions in final reports and publications. In one study, for example, 80 consecutive studies were selected for assessment of completeness from the journal Evidence-Based Medicine, a journal aimed specifically at doctors working in primary care and general medicine. Two general practitioners independently assessed whether they could use the treatment with a patient if they saw them tomorrow[3]. Of these 80 published reports, 41 (51%) of had elements of the intervention missing, particularly descriptions of process and information on hand-outs or booklets. The proportion of trials for which adequate information could be made available increased to 90% through the checking of references, contacting authors, and undertaking additional searches[3].

Similarly, Schroter et al developed, piloted and applied a checklist designed to assess the replicability of published treatment decisions to 51 trials published in the BMJ[4]. This checklist was applied by the study team to a broad range of health topics and included seven items: where the treatment was delivered (setting); who delivered the treatment (provider); who received the treatment (recipient); what was the procedure including the sequencing of the technique (procedure); a description of the physical or informational materials used (materials); the dose/duration of individual sessions of treatment (intensity); the interval, frequency, duration or timing of the treatment (schedule).This study reported that 57% (29/51) of papers were not considered to be of sufficient description to allow replication, with

the most poorly described aspects of the published trials being the sequencing of the technique (e.g. what happened and when) and physical / information materials[4].

Rates of replicability of interventions vary considerably in the published literature depending on the complexity of the treatment and the assessment criteria. For example, three studies assessed compliance with item four of CONSORT in published research in the areas of weight loss[5], brain tumours[6] and Hodgkin's lymphoma[7]. Item four of CONSORT specifically asks for precise details concerning treatments intended for all groups and how and when they were administered. These studies reported that over 90% of study findings were replicable. In contrast, however, one study assessed whether there was sufficient information on what happens before, during and after treatment for back pain, and revealed that only 13% of trials were replicable[8].

The NIHR Health Technology Assessment (HTA) Programme commissions and funds primary research and evidence synthesis on the effectiveness, costs and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the NHS. It aspires to enable all funded projects to complete and publish in the programme's own Journal '*Health Technology Assessment*', freely available on the programme's website (www.hta.ac.uk). Reports published in the Journal series are peer reviewed, are in the public domain and contain a full record of the study. Unlike typical peer reviewed journals, there are no word or size limitations for the full report and unlimited appendices, thus enabling more detail to be included in the publication. Given the importance of complete and replicable reporting of findings and the opportunities the NIHR HTA Journal presents, this study aimed to assess whether randomised controlled trials (RCTs) with single trials published in *Health Technology Assessment* were described in sufficient detail.

Methods:

An amended version of a checklist developed by Schroter et al[4] was applied to single trial projects (not projects which contained multiple trials publishing in the NIHR HTA Journal (Table 1). Schroter's original checklist assessed published studies using seven criteria. Based on the results of the pilot study we developed an twelve item checklist clustered into nine areas. This enabled separate assessments to be made of inclusion / exclusion criteria / baseline characteristics for recipient details and for physical and informational materials. We did not apply the checklist separately to the Control Group but did make a general assessment as to the completeness of Control Group information, but decided to report the findings of this separately from all other criteria as this ultimately was excluded from the Schroter checklist. The study was conducted in two phases:

***Phase 1*:**

Phase I piloted and modified a checklist initially developed by Schroter et al[4]. Five NIHR HTA funded RCTs were selected to represent a range of interventions (surgery, psychology, devices and pharmaceutical). The checklist was applied independently by three assessors to assess the level of agreement between assessments. A kappa score was produced for each individual trial (range 0.15 – 0.7) with an average of 0.225 across all 5 trials. Disagreement was due to differing interpretations of the checklist questions. The initial checklist was revised and included assessments of settings, recipients, providers, procedures, materials, intensity and schedule. The modified checklist was applied by the three assessors to a further five NIHR HTA funded RCTs, resulting in higher levels of agreement (kappa scores for each report ranged from 0.3 – 0.7, with an average of 0.6 for all trials).

Phase 2:

The revised checklist was applied to a wider sample of NIHR HTA funded RCTs. All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion with in the study. Of the 109 reports publishing in this time-period, 11 were excluded as they involved at least two trials. Ninety-eight single trial RCTs were therefore included in the analysis. The unit of analysis used for this project was one checklist assessment per trial published. Table two shows the number of journals for each intervention by type.

Data were exported from Microsoft Access to IBM SPSS version 19, and this software was used to conduct all descriptive and inferential analyses. The three team members who

carried out the assessments are all NETSCC researchers with higher health degrees. None of the researchers had medical or clinical experience. All NIHR HTA reports were examined by initially scanning the executive summary, followed by the methods, using key word search terms to scan the whole document and appendices, and finally undertaking a detailed reading of the entire journal if relevant information could not be found. Table 3 contains selected examples against six of the items in the checklist to illustrate both complete and poorly described interventions.

Disagreements in the scoring of reports were discussed in team meetings where differences in assessments were considered and resolved.

Results:

Applying the revised checklist to NIHR HTA funded RCTs revealed that components of the intervention description were missing in 68 / 98 reports (missing 69.4%) (Table 4). Intervention descriptions were therefore complete in 30.6% of reports. Certain criteria had high levels of completeness, such as baseline characteristics and descriptions of settings, which were complete for over 90% of reports. However, other criteria were notably less complete, particularly patient information with 58.2% having an adequate description (Table 4).

Differences in completion rates were noted between the types of interventions. For example, descriptions of interventions were more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively, although Chi Square analysis showed that this difference was not statistically significant (p=0.77). Furthermore, this was not the case with certain criteria, such as baseline characteristics and provider information where levels of completeness were higher in non-drug trials than drug interventions.

Descriptions of interventions were found to be least complete for psychological interventions with only 27.3% of RCTs in this area being complete, although Chi squared analysis revealed that this difference was not statistically significant when compared with drug interventions (p=1.00). Again, there were a few occasions where certain criteria had the highest levels of completeness of all intervention types, in particular with baseline characteristics and provider information with 100% and 90.9% of completeness respectively (Table 4).

Data were collected on the completeness rates for Control Group information but were not included with the full data set as this criterion was not included in Schroter's original checklist. The data revealed that 51% of RCTs had complete descriptions of control groups.

Discussion:

Statement of principle findings

This study has revealed that 30.6% (30/98) of studies with a single trial published in the HTA Journal have a full description of the intervention. The interventions described in published RCTs performed well against certain criteria, such as baseline characteristics (with 95% having an adequate description), but less well on other criteria, such as patient information (with 58% having an adequate description). Drug trials were slightly more complete than non-drug trials and psychological interventions with 33.3% of journals having a complete intervention description, although these differences were not statistically significant.

Strengths and weaknesses

The strengths of this study are that externally generated and tested criteria were applied to evaluate the effectiveness of intervention descriptions in NIHR HTA Programme funded RCTs. However, there were limitations. First, none of the assessors applying the criteria were medically trained, however, assessors were not commenting on the suitability of an intervention for use in practice but whether aspects of the description that would be required for use in practice were present. There is a possibility that someone with medical training would score the projects differently. In previous work the authors have been medically trained. Second, authors of the reports were not contacted to provide additional information beyond that provided in the publication. As Glasziou has demonstrated, contacting the research teams or additional searches for intervention details does increase the completeness of intervention descriptions[3]. However it is questionable whether having to undertake additional searches outside the publication effectively enhances the ease of replicating study findings. A third limitation concerned the type of data being collected. Whilst all the criteria are dichotomous (in that they are all yes / no answers), the justification behind this categorisation has different degrees of interpretation. This could have resulted in overly-harsh assessments of completeness for certain criteria. For example, the recipient criterion is clear compared with the greater interpretation required by the materials criterion. Certainly the completion rate for materials was among the lowest across all studies with 58% and 69% completion rates for informational and physical materials respectively.

A further limitation of the study was that a full assessment of the control group was not undertaken. The checklist could have been applied to the control group in addition to the intervention; this would have provided a more complete picture of how well controls are described within a study. Another limitation was that the number of journals assessed for completeness was very small for certain assessments (for example only 11 / 98 journals

reported psychological interventions). It is possible therefore, that certain findings of completeness rate occurred by chance.

Meanings of the study

It is tempting to make comparisons with others studies assessing the usability of intervention descriptions. In particular, Glasziou reported that 41/80 (51%) of published reports of single randomised trials and systematic reviews in popular journals were complete compared with 30/98 (30.6%) completeness of NIHR HTA funded RCT trials. Similarly, interventions in NIHR HTA reports appeared to be described less well than the 51 trials published in the BMJ assessed by Schroter et al. where 43% (22/51) of papers were considered to be of sufficient description to allow replication[4]. Whilst these comparisons are interesting, it is important to note that it is not possible to make any meaningful comparison on the relative performance of each output, as the Glasziou study looked at journal articles and we looked at the HTA journal series which are aimed at different audiences and the questionnaire used was different between the studies. This is because the nature of outputs varies considerably between studies as does the assessment criteria. It is notable, for example, that Schroter et al[4] used seven indicators in their checklist, compared with the twelve criteria used in this study.

However, this study does reflect findings from similar studies conducted elsewhere. For example, the criteria highlighted as being particularly poorly described in Schroter's study were physical / informational materials, which reflected findings in this study where patient information and physical materials were also lacking in completeness. Similarly the fact that NIHR HTA Programme funded drug interventions were typically better described than non-drug interventions reflected findings in Glasziou et al. where over 60% of reports on drug treatments were initially deemed to be complete compared with just under 30% of non- drug treatments.

Understanding the extent to which interventions in published studies are described sufficiently to inform clinical decision making is a key concern in the adding value in research agenda. As Chalmers and Glasziou have suggested, poorly described interventions form one of the four main pillars of research waste[1]. The criteria identified by Schroter et al[4] and developed in this study are helpful in highlighting specific areas of where intervention descriptions can be improved.

Future research

Several areas for further research are indicated by this study. Further testing on the criteria can be undertaken to assess the repeatability of the criteria. For example, the reports sampled in this study could be reassessed by someone with clinical experience to assess the level of agreement. Alternatively, Glasziou's selected papers in his original study could be assessed by non-clinical teams to examine the level of agreement.

Ensuring the replicability of study findings is an essential part of adding value in research. It is important for health research publishers to be transparent in the usability of study reports and areas of improvement. This study applied a checklist that can be used to indicate where the descriptions of interventions can be improved to enhance replication in clinical practice. Serious consideration should be given on how this might be used to improve intervention reporting in the future.

References

1. Chalmers I,Glasziou P. Avoidable waste in the production and reporting of research evidence. *The Lancet* 2009;**374**:86-9
2. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ* 2010; 341.
3. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008; 336(7659):1472-1474.
4. Schroter S, Glasziou P, Heneghan C. Quality of description of treatments: a review of published randomised controlled trials. BMJ Open 2012; 2.
5. Thabane L, Chu R, Cuddy K, Douketis J. What is the quality of reporting in weight loss intervention studies? A systematic review of randomized controlled trials. [Review] [17 refs]. *International Journal of Obesity* 2007; 31(10):1554-1559
6. Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in Hodgkin's lymphoma in biomedial journals. *J Natl Cancer Inst* 2006; 98:620-625.
7. Lai R, Chu R, Fraumeni M, et al. Quality of ransomized controlled trials reporting in the primary treamt of brain tumors. *J Clin Oncol* 2006; 24:1136-1144.
8. Glenton C, Underland V, Kho M, Pennick V, Oxman AD. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. *Journal of Clinical Epidemiology* 2006; 59(8):770-778

Table 1: Replicability Criteria for Interventions, developed from an initial design by Schroter et al[3]

Checklist criteria		Descriptor of criteria where appropriate
1. Setting	Is it clear where the intervention was delivered?	
2a. Recipient – inclusion	Is it clear who is receiving the intervention? – Inclusion criteria	Clear inclusion criteria in the journal
2b. Recipient – exclusion	Is it clear who is receiving the intervention? – Exclusion criteria	Clear exclusion criteria in journal
2c. Recipient – baseline characteristics	Do you know all that you need to about the patients? (e.g. which drugs they are taking, what they were told, etc)? If No, what further information do you require?	Baseline characteristics of participants provided in journal
3. Provider	Is it clear who delivered the intervention?	
4. Procedure	Is the procedure (including the sequencing of the technique) of the intervention sufficiently clear to allow replication?	Top level overview of the intervention. Eg Drug X for X days at X dose. Or X sessions lasting X minutes, for X weeks/ months
5. Intensity	Is the dose/duration of individual sessions of the intervention clear?	Dose, length of session
6. Schedule	Is the schedule (interval, frequency, duration, or timing) of the intervention clear?	Frequency of intervention, length of session
7a. Materials – physical	Are the physical materials used adequately described?	Physical materials eg Description of splint used. If either are no, it is a no overall
7b. Materials - informational	Are the informational materials used adequately described?	Information provided to the patients eg consent forms etc
8. Missing	Is the description of the intervention complete? If No, what is missing?	
9. Control	Is it clear what the control group received during the study?	

Table 2: Intervention type of NIHR HTA funded trial RCTs included in the study

Type of intervention	№ (%)
Drug	15 (15.3)
Radiotherapy	1 (1.0)
Surgery	9 (9.2)
Diagnostic	8 (8.2)
Education and training	3 (3.1)
Service delivery	19 (19.4)
Psychological therapies	11 (11.2)
Vaccines and biologicals	3 (3.1)
Devices	12 (12.2)
Physical therapies	7 (7.1)
Exercise	1 (1.0)
Complementary therapies	2 (2.0)
Mixed or complex	6 (6.1)
Other*	1 (1.0)
Total	**98**

*Other refers to an intervention using larval therapy

Table 3: Examples of poor reporting of intervention elements within the HTA journal series, taken verbatim form the journal.

Checklist item	Examples of poor reporting	Reason why rated as incomplete	Examples of good reporting	Reason why rated complete
Inclusion criteria	'….patient identification was retrospective. Searches were conducted on practice databases using either repeat prescriptions alone or repeat prescriptions plus diagnostic terms… GPs then sent letters to suitable patients, providing information about the trial'	No details given about the searches and the criteria patients were screened with.	'Inclusion criteria for trial patients were: -Diagnosed with idiopathic arthritides of childhood with onset before their 16th birthday for more than 3 months. - Aged 4–19 years inclusive. - Stable on medication. - At least one active joint, core set criteria 1.56 -At least two out of any five of the remaining core set criteria below. -The physician global assessment of disease activity >10 mm on a 100-mm visual analogue scale (VAS). - The parent global assessment of well-being >10 mm on a 100-mm VAS. -Childhood Health Assessment Questionnaire scores >0. - More than one joint with limited range of motion (joint motion reduced by at least 5° from normative range for age58). -An elevated erythrocyte sedimentation rate (ESR) (>5 mmHg in children and >10 mmHg in adolescents)'	Very detailed patient criteria listed
Exclusion criteria	'GPs were given a ringbinder file with information and instructions about the trial and, within each, a number of recruitment packs. The packs contained the paperwork required to complete the recruitment of each patient, this was: a reminder of the inclusion/exclusion criteria for the study….'	No details given about the exclusion criteria	'Reasons for Exclusion (Yes/No) - BMI > 40 kg/m2 - Barrett's oesophagus (≥3cm) - Paraoesophageal hernia - Oesophageal strictures - One type of management is clinically indicated for another reason'	Detailed patient exclusion criteria listed
Provider	'All services had staff who were trained and experienced in family therapy, but not necessarily family interventions specifically for eating disorders'	No details about the staff providing the interventions or the training they received.	'Eight counsellors (six females and two males) took part in the trial (one worked at two practices) and all were BAC accredited or eligible for BAC accreditation; they were highly trained and had considerable experience of counselling in a general practice setting' (There are details about each counsellors age, qualifications and experience are provided)	States who delivered the intervention and their training
Procedure	'Generally home-based rehabilitation services provide, as a minimum, physiotherapy and occupational therapy in the patient's own home. Services can be specialised (e.g. in stroke rehabilitation) or be provided for patients with a range of disabilities'	No details about the services provided to patients and variation between centres	'The content of the CBT programme included (complete course description contained within an appendix): - Elucidation of core beliefs regarding their illness and its management. - Monitoring of activity levels and introduction of appropriate timetable. - Introduction to exercises designed to increase general level of fitness, balance and confidence in exercise. A range of aerobic, strength, balance and stretching exercises were taught. - Behavioural modification of sleep patterns. - Mood management advice. - Goal setting'	Key aspects of the intervention summarised in the text and a full description of the intervention is detailed in the appendices.
Intensity and Schedule	'Patients come to the day hospital where the rehabilitation service is provided for a full or half day. Usually ambulance transport is provided to bring patients into the service and return them home after a session'	No details of the length or number of sessions	'Psychological treatment was based on existing protocols (references included) and distributed over six 50-minute sessions, with printed information sheets provided after each session'	The length and number of sessions is included as well as the details of each session.
Materials – physical	'The acupuncture point prescriptions used were individualised to each patient and were at the discretion of the acupuncturist'	The prescriptions used are not detailed.	'- 500 mg oral oxytetracycline (non-proprietary) b.d. + topical vehicle control b.d. -100 mg oral Minocin MR minocycline) o.d. + topical vehicle control b.d. -topical Panoxyl Aquagel (5% benzoyl peroxide) b.d. + oral placebo o.d. This was designated as the active comparator group, as benzoyl peroxide was the leading and most established topical treatment for acne when the protocol was written. -topical Benzamycin (3% erythromycin + 5% benzoyl peroxide) b.d. + oral placebo o.d. (referred to as ery. + BP bd) -topical Stiemycin (2% erythromycin) o.d. +topical Panoxyl Aquagel (5% benzoyl peroxide)o.d. + oral placebo o.d. (referred to as ery. od+ BP od)'	Each of the treatments prescribed is clearly defined

Table 4: Completion rates of NIHR HTA reports by intervention type

Description criteria	Drugs No. Trials with complete intervention description (%)	CI*	Non-Drugs No. Trials with complete intervention description (%)	CI*	Psychological No. Trials with complete intervention description (%)	CI*	All No. Trials with complete intervention description (%)
Number journals	15 (15.3)		72 (73.5)		11 (11.2)		98 (100)
Setting	15 (100)	0.80-1	66 (91.7)	0.83-0.96	9 (81.8)	0.52-0.95	90 (91.8)
Inclusion criteria	15 (100)	0.80-1	63 (87.5)	0.78-0.93	10 (90.9)	0.62-0.98	88 (89.8)
Exclusion criteria	14 (93.3)	0.70-0.99	55 (76.4)	0.65-0.85	10 (90.9)	0.62-0.98	79 (80.6)
Baseline characteristics	14 (93.3)	0.70-0.99	68 (94.4)	0.87-0.98	11 (100)	0.74-1	93 (94.9)
Provider	11 (73.3)	0.48-0.89	56 (77.8)	0.67-0.86	10 (90.9)	0.62-0.98	77 (78.6)
Procedure	14 (93.3)	0.70-0.99	57 (79.2)	0.68-0.87	9 (81.8)	0.52-0.95	80 (81.6)
Intensity	13 (86.7)	0.62-0.96	63 (87.5)	0.78-0.93	9 (81.8)	0.52-0.95	85 (86.8)
Schedule	13 (86.7)	0.62-0.96	59 (81.9)	0.71-0.89	9 (81.8)	0.52-0.95	81 (82.7)
Patient information	9 (60.0)	0.36-0.80	42 (58.3)	0.47-0.69	6 (54.5)	0.28-0.79	57 (58.2)
Physical materials	10 (66.7)	0.42-0.85	52 (72.3)	0.61-0.81	6 (54.5)	0.28-0.79	68 (69.4)
Intervention description complete overall	5 (33.3)	0.15-0.58	22 (30.6)	0.21-0.42	3 (27.3)	0.10-0.57	30 (30.6)

**Please note: some criteria are not applicable therefore denominator less than total number of journals.*95% confidence interval: no continuity correction*

The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-003713.R1
Article Type:	Research
Date Submitted by the Author:	18-Oct-2013
Complete List of Authors:	Douet, Lisa; University of Southampton, National Institute for Health Research Milne, Ruairidh; University of Southampton, Wessex Institute Anstee, Sydney; University of Southampton, National Institute for Health Research Habens, Fay; University of Southampton, National Institute for Health Research Young, Amanda; University of Southampton, NETSCC Wright, David; University of Southampton, National Institute for Health Research
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Research methods
Keywords:	Clinical trials < THERAPEUTICS, AUDIT, STATISTICS & RESEARCH METHODS

The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Dr Lisa Douet[a]
Senior Research Fellow

Professor Ruairidh Milne[b]
Professorial Fellow in Public Health

Sydney Anstee[a]
Research Fellow

Fay Habens[a]
Programme Manager

Amanda Young[a]
Senior Research Fellow

Dr David Wright[a]
Senior Research Fellow

[a]National Institute for Health Research
Evaluation, Trials and Studies Coordinating Centre (NETSCC)
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

[b]Wessex Institute
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

Correspondence to: l.douet@southampton.ac.uk

Abstract word count: 239
Main text word count: 2480

Abstract:

Objectives

The objective of this study is to assess whether NIHR HTA funded randomised controlled trials (RCTs) published in *Health Technology Assessment* Journal were described in sufficient detail to replicate in practice.

Methods

A published checklist for assessing intervention descriptions was applied to NIHR HTA funded RCTs published in *Health Technology Assessment*. The checklist was piloted twice on a sample of 10 reports. Kappa scores were generated to assess agreement in the checklist application. The checklist was modified and applied to all 98 NIHR HTA funded single trial RCTs published in the Journal up to March 2011. Three assessors independently applied the checklist. Disagreements in scoring were discussed in the team; differences were then explored and resolved.

Results

Components of the intervention description were missing in 68 / 98 (69.4%) reports. Baseline characteristics and descriptions of settings had the highest levels of completeness with over 90% of reports complete. Reports were less complete on patient information with 58.2% of the journals having an adequate description. Intervention descriptions were more not significantly more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively. Only 27.3% of RCTs with psychological interventions were deemed to be complete, although numbers were too small for differences to be significant statistically.

Conclusions

Ensuring the replicability of study interventions is an essential part of adding value in research. Research funders need to ensure transparency and completeness in the reporting of interventions.

Article summary

Article focus

- It has previously been suggested that over 50% of intervention descriptions are not sufficiently described. This article investigates the description of interventions within the NIHR HTA journal series.

Key messages

- Only 30.6% (30/98) of studies with a single trial published in the NIHR HTA Journal have a full description of the intervention
- The unlimited word count of the NIHR journal series does not affect completeness of an intervention description

Strengths and limitations

- An externally produced checklist was applied to all RCTs publishing in the NIHR HTA monograph

Introduction:

A recent publication by Glasziou and Chalmers has suggested that as much as 85% of the US$100 billion spent on health research worldwide each year is potentially wasted due to four key problems of knowledge production and dissemination. These four areas include: 1) ensuring the right research questions are asked; 2) ensuring that study designs are appropriate and are of methodological quality; 3) ensuring the findings from funded research are available in the public domain; 4) ensuring that funded research is unbiased and usable[1].

Several studies have specifically assessed the fourth area of waste by exploring the quality and usability of publications from funded health research. This is a key concern considering the role effective summaries of evidence have in facilitating knowledge transfer and enhancing the uptake of findings in clinical practice. Whilst it is recognised that trial registration databases and scientific journals are restrictive in terms of word allowance, various strategies have been proposed to improve the reporting of interventions in published trials, including an 'intervention bank' to include manuals and fidelity tools linked to trial registration numbers'[2].

Studies have highlighted concerns about the descriptions of interventions in final reports and publications. In one study, for example, 80 consecutive studies were selected for assessment of completeness from the journal Evidence-Based Medicine, a journal aimed specifically at doctors working in primary care and general medicine. Two general practitioners independently assessed whether they could use the treatment with a patient if they saw them tomorrow[3]. Of these 80 published reports, 41 (51%) of had elements of the intervention missing, particularly descriptions of process and information on hand-outs or booklets. The proportion of trials for which adequate information could be made available increased to 90% through the checking of references, contacting authors, and undertaking additional searches[3].

Similarly, Schroter et al developed, piloted and applied a checklist designed to assess the replicability of published treatment decisions to 51 trials published in the BMJ[4]. This checklist was applied by the study team to a broad range of health topics and included seven items: where the treatment was delivered (setting); who delivered the treatment (provider); who received the treatment (recipient); what was the procedure including the sequencing of the technique (procedure); a description of the physical or informational materials used (materials); the dose/duration of individual sessions of treatment (intensity); the interval, frequency, duration or timing of the treatment (schedule).This study reported that 57% (29/51) of papers were not considered to be of sufficient description to allow replication, with

the most poorly described aspects of the published trials being the sequencing of the technique (e.g. what happened and when) and physical / information materials[4].

Rates of replicability of interventions vary considerably in the published literature depending on the complexity of the treatment and the assessment criteria. For example, three studies assessed compliance with item four of CONSORT in published research in the areas of weight loss[5], brain tumours[6] and Hodgkin's lymphoma[7]. Item four of CONSORT specifically asks for precise details concerning treatments intended for all groups and how and when they were administered. These studies reported that over 90% of study findings were replicable. In contrast, however, one study assessed whether there was sufficient information on what happens before, during and after treatment for back pain, and revealed that only 13% of trials were replicable[8].

The NIHR Health Technology Assessment (HTA) Programme commissions and funds primary research and evidence synthesis on the effectiveness, costs and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the NHS. It aspires to enable all funded projects to complete and publish in the programme's own Journal '*Health Technology Assessment*', freely available on the programme's website (www.hta.ac.uk). Reports published in the Journal series are peer reviewed, are in the public domain and contain a full record of the study. Unlike typical peer reviewed journals, there are no word or size limitations for the full report and unlimited appendices, thus enabling more detail to be included in the publication. Given the importance of complete and replicable reporting of findings and the opportunities the NIHR HTA Journal presents, this study aimed to assess whether randomised controlled trials (RCTs) with single trials published in *Health Technology Assessment* were described in sufficient detail.

Methods:

An amended version of a checklist developed by Schroter et al[4] was applied to single trial projects (not projects which contained multiple trials publishing in the NIHR HTA Journal (Table 1). Schroter's original checklist assessed published studies using seven criteria. Based on the results of the pilot study we developed an twelve item checklist clustered into nine areas. This enabled separate assessments to be made of inclusion / exclusion criteria / baseline characteristics for recipient details and for physical and informational materials. We did not apply the checklist separately to the Control Group but did make a general assessment as to the completeness of Control Group information, but decided to report the findings of this separately from all other criteria as this ultimately was excluded from the Schroter checklist. The study was conducted in two phases:

***Phase 1*:**

Phase I piloted and modified a checklist initially developed by Schroter et al[4]. Five NIHR HTA funded RCTs were selected to represent a range of interventions (surgery, psychology, devices and pharmaceutical). The checklist was applied independently by three assessors to assess the level of agreement between assessments. A kappa score was produced for each individual trial (range 0.15 – 0.7) with an average of 0.225 across all 5 trials. Disagreement was due to differing interpretations of the checklist questions. The initial checklist was revised and included assessments of settings, recipients, providers, procedures, materials, intensity and schedule. The modified checklist was applied by the three assessors to a further five NIHR HTA funded RCTs, resulting in higher levels of agreement (kappa scores for each report ranged from 0.3 – 0.7, with an average of 0.6 for all trials).

Phase 2:

The revised checklist was applied to a wider sample of NIHR HTA funded RCTs. All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion with in the study. Of the 109 reports publishing in this time-period, 11 were excluded as they involved at least two trials. Ninety-eight single trial RCTs were therefore included in the analysis. The unit of analysis used for this project was one checklist assessment per trial published. Table two shows the number of journals for each intervention by type.

Data were exported from Microsoft Access to IBM SPSS version 19, and this software was used to conduct all descriptive and inferential analyses. The three team members who

carried out the assessments are all NETSCC researchers with higher health degrees. None of the researchers had medical or clinical experience. All NIHR HTA reports were examined by initially scanning the executive summary, followed by the methods, using key word search terms to scan the whole document and appendices, and finally undertaking a detailed reading of the entire journal if relevant information could not be found. Table 3 contains selected examples against six of the items in the checklist to illustrate both complete and poorly described interventions.

Disagreements in the scoring of reports were discussed in team meetings where differences in assessments were considered and resolved.

Results:

Applying the revised checklist to NIHR HTA funded RCTs revealed that components of the intervention description were missing in 68 / 98 reports (missing 69.4%) (Table 4). Intervention descriptions were therefore complete in 30.6% of reports. Certain criteria had high levels of completeness, such as baseline characteristics and descriptions of settings, which were complete for over 90% of reports. However, other criteria were notably less complete, particularly patient information with 58.2% having an adequate description (Table 4).

Differences in completion rates were noted between the types of interventions. For example, descriptions of interventions were more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively, although Chi Square analysis showed that this difference was not statistically significant (p=0.77). Furthermore, this was not the case with certain criteria, such as baseline characteristics and provider information where levels of completeness were higher in non-drug trials than drug interventions.

Descriptions of interventions were found to be least complete for psychological interventions with only 27.3% of RCTs in this area being complete, although Chi squared analysis revealed that this difference was not statistically significant when compared with drug interventions (p=1.00). Again, there were a few occasions where certain criteria had the highest levels of completeness of all intervention types, in particular with baseline characteristics and provider information with 100% and 90.9% of completeness respectively (Table 4).

Data were collected on the completeness rates for Control Group information but were not included with the full data set as this criterion was not included in Schroter's original checklist. The data revealed that 51% of RCTs had complete descriptions of control groups.

Discussion:

Statement of principle findings

This study has revealed that 30.6% (30/98) of studies with a single trial published in the HTA Journal have a full description of the intervention. The interventions described in published RCTs performed well against certain criteria, such as baseline characteristics (with 95% having an adequate description), but less well on other criteria, such as patient information (with 58% having an adequate description). Drug trials were slightly more complete than non-drug trials and psychological interventions with 33.3% of journals having a complete intervention description, although these differences were not statistically significant.

Strengths and weaknesses

The strengths of this study are that externally generated and tested criteria were applied to evaluate the effectiveness of intervention descriptions in NIHR HTA Programme funded RCTs. However, there were limitations. First, none of the assessors applying the criteria were medically trained, however, assessors were not commenting on the suitability of an intervention for use in practice but whether aspects of the description that would be required for use in practice were present. There is a possibility that someone with medical training would score the projects differently. In previous work the authors have been medically trained. Second, authors of the reports were not contacted to provide additional information beyond that provided in the publication. As Glasziou has demonstrated, contacting the research teams or additional searches for intervention details does increase the completeness of intervention descriptions[3]. However it is questionable whether having to undertake additional searches outside the publication effectively enhances the ease of replicating study findings. A third limitation concerned the type of data being collected. Whilst all the criteria are dichotomous (in that they are all yes / no answers), the justification behind this categorisation has different degrees of interpretation. This could have resulted in overly-harsh assessments of completeness for certain criteria. For example, the recipient criterion is clear compared with the greater interpretation required by the materials criterion. Certainly the completion rate for materials was among the lowest across all studies with 58% and 69% completion rates for informational and physical materials respectively.

A further limitation of the study was that a full assessment of the control group was not undertaken. The checklist could have been applied to the control group in addition to the intervention; this would have provided a more complete picture of how well controls are described within a study. Another limitation was that the number of journals assessed for completeness was very small for certain assessments (for example only 11 / 98 journals

reported psychological interventions). It is possible therefore, that certain findings of completeness rate occurred by chance.

Meanings of the study

It is tempting to make comparisons with others studies assessing the usability of intervention descriptions. In particular, Glasziou reported that 41/80 (51%) of published reports of single randomised trials and systematic reviews in popular journals were complete compared with 30/98 (30.6%) completeness of NIHR HTA funded RCT trials. Similarly, interventions in NIHR HTA reports appeared to be described less well than the 51 trials published in the BMJ assessed by Schroter et al. where 43% (22/51) of papers were considered to be of sufficient description to allow replication[4]. Whilst these comparisons are interesting, it is important to note that it is not possible to make any meaningful comparison on the relative performance of each output, as the Glasziou study looked at journal articles and we looked at the HTA journal series which are aimed at different audiences and the questionnaire used was different between the studies. This is because the nature of outputs varies considerably between studies as does the assessment criteria. It is notable, for example, that Schroter et al[4] used seven indicators in their checklist, compared with the twelve criteria used in this study.

However, this study does reflect findings from similar studies conducted elsewhere. For example, the criteria highlighted as being particularly poorly described in Schroter's study were physical / informational materials, which reflected findings in this study where patient information and physical materials were also lacking in completeness. Similarly the fact that NIHR HTA Programme funded drug interventions were typically better described than non-drug interventions reflected findings in Glasziou et al. where over 60% of reports on drug treatments were initially deemed to be complete compared with just under 30% of non- drug treatments.

Understanding the extent to which interventions in published studies are described sufficiently to inform clinical decision making is a key concern in the adding value in research agenda. As Chalmers and Glasziou have suggested, poorly described interventions form one of the four main pillars of research waste[1]. The criteria identified by Schroter et al[4] and developed in this study are helpful in highlighting specific areas of where intervention descriptions can be improved.

Future research

Several areas for further research are indicated by this study. Further testing on the criteria can be undertaken to assess the repeatability of the criteria. For example, the reports sampled in this study could be reassessed by someone with clinical experience to assess the level of agreement. Alternatively, Glasziou's selected papers in his original study could be assessed by non-clinical teams to examine the level of agreement.

Ensuring the replicability of study findings is an essential part of adding value in research. It is important for health research publishers to be transparent in the usability of study reports and areas of improvement. This study applied a checklist that can be used to indicate where the descriptions of interventions can be improved to enhance replication in clinical practice. Serious consideration should be given on how this might be used to improve intervention reporting in the future.

Acknowledgements:
We would like to acknowledge Professor James Raftery and the Metadata team for providing the database and the trial details used in the study and for Professor Paul Glasziou for his advice during the study.

Ethic statement:
This study did not require ethics approval as no data patient or clinical data was required.

Funding statement:
This study was supported by the NIHR Evaluation, Trials and Studies Coordinating Centre through its Research on Research Programme. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Department of Health, or of NETSCC.

Contribution of Authors:
The study was conceived and designed by LD, RM, SA, and FH, and undertaken by LD, SA and FH; AY and DW supported the data analysis. All authors read and approved the final manuscript. LD is guarantor of the study

Data sharing: Data on the included trials are available on request from the corresponding author.

Competing Interests: The authors have no competing financial interests; however, all of the authors are employed by the University of Southampton to work at least part time for NETSCC. In particular: RM is employed as the Head of NETSCC and has worked for NETSCC (and its predecessor organisation) in senior roles on and off since 1996. He was an editor of the Health Technology Assessment journal (1997-2007) and a founder editor for other journals in the new NIHR Journals Library (2011-12).

References

1. Chalmers I,Glasziou P. Avoidable waste in the production and reporting of research evidence. *The Lancet* 2009;**374**:86-9
2. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ* 2010; 341.
3. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008; 336(7659):1472-1474.
4. Schroter S, Glasziou P, Heneghan C. Quality of description of treatments: a review of published randomised controlled trials. BMJ Open 2012; 2.
5. Thabane L, Chu R, Cuddy K, Douketis J. What is the quality of reporting in weight loss intervention studies? A systematic review of randomized controlled trials. [Review] [17 refs]. *International Journal of Obesity* 2007; 31(10):1554-1559
6. Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in Hodgkin's lymphoma in biomedial journals. *J Natl Cancer Inst* 2006; 98:620-625.
7. Lai R, Chu R, Fraumeni M, et al. Quality of ransomized controlled trials reporting in the primary treamt of brain tumors. *J Clin Oncol* 2006; 24:1136-1144.
8. Glenton C, Underland V, Kho M, Pennick V, Oxman AD. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. *Journal of Clinical Epidemiology* 2006; 59(8):770-778

The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Dr Lisa Douet[a]
Senior Research Fellow

Professor Ruairidh Milne[b]
Professorial Fellow in Public Health

Sydney Anstee[a]
Research Fellow

Fay Habens[a]
Programme Manager

Amanda Young[a]
Senior Research Fellow

Dr David Wright[a]
Senior Research Fellow

[a]National Institute for Health Research
Evaluation, Trials and Studies Coordinating Centre (NETSCC)
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

[b]Wessex Institute
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

Correspondence to: l.douet@southampton.ac.uk

Acknowledgements:
We would like to acknowledge Professor James Raftery and the Metadata team for providing the database and the trial details used in the study and for Professor Paul Glasziou for his advice during the study.

Ethic statement:
This study did not require ethics approval as no data patient or clinical data was required.

Funding statement:
This study was supported by the NIHR Evaluation, Trials and Studies Coordinating Centre through its Research on Research Programme. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Department of Health, or of NETSCC.

Contribution of Authors:
The study was conceived and designed by LD, RM, SA, and FH, and undertaken by LD, SA and FH; AY and DW supported the data analysis. All authors read and approved the final manuscript. LD is guarantor of the study

Data sharing: Data on the included trials are available on request from the corresponding author.

Competing Interests: The authors have no competing financial interests; however, all of the authors are employed by the University of Southampton to work at least part time for NETSCC. In particular: RM is employed as the Head of NETSCC and has worked for NETSCC (and its predecessor organisation) in senior roles on and off since 1996. He was an editor of the Health Technology Assessment journal (1997-2007) and a founder editor for other journals in the new NIHR Journals Library (2011-12).

Abstract word count: 239
Main text word count: 2480

Abstract:

Objectives

The objective of this study is to assess whether NIHR HTA funded randomised controlled trials (RCTs) published in *Health Technology Assessment* Journal were described in sufficient detail to replicate in practice.

Methods

A published checklist for assessing intervention descriptions was applied to NIHR HTA funded RCTs published in *Health Technology Assessment*. The checklist was piloted twice on a sample of 10 reports and modified. Kappa scores were generated to assess agreement in the checklist application. The modified checklist was ~~modified and~~ applied to all 98 NIHR HTA funded single trial RCTs published in the Journal up to March 2011. The checklist included assessments of participants, intensity, schedule, materials and settings. A study was agreed to be complete overall, is aspects of the checklist were present. ~~Three a~~Assessors independently applied the checklist. Disagreements in scoring were discussed in the team; differences were then explored and resolved.

Results

Components of the intervention description were missing in 68 / 98 (69.4%) reports. Baseline characteristics and descriptions of settings had the highest levels of completeness with over 90% of reports complete. Reports were less complete on patient information with 58.2% of the journals having an adequate description. When looking at individual intervention types, drug intervention descriptions were more complete than non-drug interventions with 33.3% and 30.6% levels of completeness respectively, although this was not significant statistically. ~~Intervention descriptions were more not significantly more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively.~~ Only 27.3% of RCTs with psychological interventions were deemed to be complete, although ~~numbers were too small for~~again these differences ~~to be~~were not significant statistically.

Conclusions

Ensuring the replicability of study interventions is an essential part of adding value in research. ~~Research~~ All those publishing clinical trial data ~~funders~~ need to ensure transparency and completeness in the reporting of interventions to ensure that ~~study~~ interventions can be replicated.

Article summary

Article focus

- ~~It has previously been suggested that over 50% of intervention descriptions are not sufficiently described.~~ This article investigates the adequacy of reporting of different elements of ~~description of~~ interventions reported~~s~~ in randomised controlled trials published within ~~within~~ the NIHR HTA journal series, using a simple checklist~~.~~

Key messages

- Only 30.6% (30/98) of studies with a single trial published in the NIHR HTA Journal have a full description of the intervention
- The unlimited word count of the NIHR journal series does not affect completeness of an intervention description

Strengths and limitations

- An externally produced checklist was applied to all RCTs publishing in the NIHR HTA ~~monograph~~Journal Series

Introduction:

A recent publication by Glasziou and Chalmers has suggested that as much as 85% of the US$100 billion spent on health research worldwide each year is potentially wasted due to four key problems of knowledge production and dissemination~~. These four areas include: 1) ensuring the right research questions are asked; 2) ensuring that study designs are appropriate and are of methodological quality; 3) ensuring the findings from funded research are available in the public domain; 4) ensuring that funded research is unbiased and usable~~[1].

Several studies have specifically assessed the waste ~~fourth area of~~area of ensuring funded research is unbiased and usable ~~waste~~ by exploring the quality and usability of publications from funded health research. This is a key concern considering the role effective summaries of evidence have in facilitating knowledge transfer and enhancing the uptake of findings in clinical practice. Whilst it is recognised that trial registration databases and scientific journals ~~are~~ can be restrictive in terms of word allowance, various strategies have been proposed to improve the reporting of interventions in published trials, including an 'intervention bank' to include manuals and fidelity tools linked to trial registration numbers'[2].

Studies have highlighted concerns about the descriptions of interventions in final reports and publications. In one study, for example, 80 consecutive studies were selected for assessment of completeness from the journal Evidence-Based Medicine, a journal aimed specifically at doctors working in primary care and general medicine. Two general practitioners independently assessed whether they could use the treatment with a patient if they saw them tomorrow[3]. Of these 80 published reports, 41 (51%) of had elements of the intervention missing, particularly descriptions of process and information on hand-outs or booklets. The proportion of trials for which adequate information could be made available increased to 90% through the checking of references, contacting authors, and undertaking additional searches[3].

Similarly, *Schroter et al* developed, piloted and applied a checklist designed to assess the replicability of published treatment decisions to 51 trials published in the BMJ[4]. This checklist was applied by the study team to a broad range of health topics and included seven items and a global eighth item to summarise completeness. ~~: where the treatment was delivered (setting); who delivered the treatment (provider); who received the treatment (recipient); what was the procedure including the sequencing of the technique (procedure); a description of the physical or informational materials used (materials); the dose/duration of individual sessions of treatment (intensity); the interval, frequency, duration or timing of the treatment (schedule).~~This study reported that 57% (29/51) of papers were not considered to be of

Formatted: Font: Italic

sufficient description to allow replication, with the most poorly described aspects of the published trials being the sequencing of the technique ~~(e.g. what happened and when)~~ and physical / information materials[4]. A further study[9] has used the checklist developed by *Schroter et al* to assess the completeness of non-pharmacological intervention description and reported that only 39% were adequately described.

Formatted: Font: Italic

Formatted: Font: Not Italic

Rates of replicability of interventions vary considerably in the published literature depending on the complexity of the treatment and the assessment criteria. For example, three studies assessed compliance with item four of CONSORT in published research in the areas of weight loss[5], brain tumours[6] and Hodgkin's lymphoma[7]. Item four of CONSORT specifically asks for precise details concerning treatments intended for all groups and how and when they were administered. These studies reported that over 90% of study findings were replicable. In contrast, however, one study assessed whether there was sufficient information on what happens before, during and after treatment for back pain, and revealed that only 13% of trials were replicable[8].

The NIHR Health Technology Assessment (HTA) Programme commissions and funds primary research and evidence synthesis on the effectiveness, costs and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the NHS. It aspires to enable all funded projects to complete and publish in the programme's own Journal '*Health Technology Assessment*', freely available on the programme's website (www.hta.ac.uk). Reports published in the Journal series are peer reviewed, are in the public domain and contain a full record of the study. Unlike typical peer reviewed journals, there are no word or size limitations for the full report and unlimited appendices, thus enabling more detail to be included in the ~~publication~~publication; an average report is approximately 50,000 in length. ~~.~~ Given the importance of complete and replicable reporting of findings and the opportunities the NIHR HTA Journal presents, this study aimed to assess whether randomised controlled trials (RCTs) with single trials published in *Health Technology Assessment* were described in sufficient detail.

Methods:

Data source

All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion in the study. Of the 109 reports publishing in this time-period, 11 were excluded as they reported more than one RCT within a single HTA Journal. Ninety-eight single trial RCTs were therefore included in the study.

~~An amended version of a checklist developed by Schroter et al[4] was applied to single trial projects (not projects which contained multiple trials publishing in the NIHR HTA Journal (Table 1). Schroter's original checklist assessed published studies using seven criteria. Based on the results of the pilot study we developed an twelve item checklist clustered into nine areas. This enabled separate assessments to be made of inclusion / exclusion criteria / baseline characteristics for recipient details and for physical and informational materials. We did not apply the checklist separately to the Control Group but did make a general assessment as to the completeness of Control Group information, but decided to report the findings of this separately from all other criteria as this ultimately was excluded from the Schroter checklist. The study was conducted in two phases:~~

Piloting the checklist

Five NIHR HTA funded RCTs were selected to represent a range of interventions (surgery, psychology, devices and pharmaceutical) to pilot the checklist initially developed by *Schroter et al*[4]. The checklist was applied independently by three assessors to assess the level of agreement between assessments. A kappa score was produced for each individual trial (range 0.15 – 0.7) with an average of 0.225 across all 5 trials. Disagreement was due to differing interpretations of the checklist questions. The initial checklist was modified to separate out two of the questions into their individual components. In the published checklist the recipient question stated 'Is it clear who is receiving the intervention?' and 'Do you know all that you need to about the patients? (e.g. which drugs they are taking, what they were told, etc)?'. We felt that this required several pieces of information for a single question and therefore we separated question 2 into the three components, as shown in table 1. Similarly the materials question, 'Are the physical or informational materials used adequately described?' was separated into the two components shown in question 7 of table 1. Additionally, the assessors discussed the type and level of information expected to be present in order to answer a question as complete. The modified checklist was applied by the three assessors to a further five NIHR HTA funded RCTs, resulting in higher levels of agreement (kappa scores for each report ranged from 0.3 – 0.7, with an average of 0.6 for all trials).

~~***Phase 1:***~~

Formatted: Font: Italic

~~Phase I piloted and modified a checklist initially developed by Schroter et al[4]. Five NIHR HTA funded RCTs were selected to represent a range of interventions (surgery, psychology, devices and pharmaceutical). The checklist was applied independently by three assessors to assess the level of agreement between assessments. A kappa score was produced for each individual trial (range 0.15 – 0.7) with an average of 0.225 across all 5 trials. Disagreement was due to differing interpretations of the checklist questions. The initial checklist was revised and included assessments of settings, recipients, providers, procedures, materials, intensity and schedule. The modified checklist was applied by the three assessors to a further five NIHR HTA funded RCTs, resulting in higher levels of agreement (kappa scores for each report ranged from 0.3 – 0.7, with an average of 0.6 for all trials).~~

~~Phase 2:~~The main study

The final ~~revised~~ modified checklist was applied to a wider sample of NIHR HTA funded RCTs. All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion with in the study. ~~Of the 109 reports publishing in this time period, 11 were excluded as they involved at least two trials. Ninety eight single trial RCTs were therefore included in the analysis. The unit of analysis used for this project was o~~One checklist was completed for the intervention group of each trial published~~assessment per trial published~~. ~~-~~ Each item in the checklist was answered by either a yes, no or not applicable response. We did not apply the checklist separately to the control group but did make a general assessment as to the completeness of control group information within the question nine of the checklist. This question was in the original checklist provided to us but was ultimately was excluded from the published *Schroter et al*[4] checklist. However, responses to this question were not on a detailed assessment of all components of the control group, unlike the intervention group itself. Question eight summarises whether there are any aspects of the intervention missing based on the responses to the previous seven questions.

Data quality

Each trial was assessed independently by two assessors. 15.% of published reports (15/98) were discussed due to disagreements of the scoring mainly around checklist item seven. All ~~D~~disagreements ~~in the scoring of reports~~ were discussed ~~in team meetings where differences in assessments~~by the team and were resolved by ~~were considered and resolved~~consensus. ~~-~~

~~Table two shows the number of journals for each intervention by type.~~

Formatted: Font: Italic

Formatted: Font: +Body CS (Arial)

Formatted: Plain Text, Justified

~~Data were exported from Microsoft Access to IBM SPSS version 19, and this software was used to conduct all descriptive and inferential analyses.~~ ~~The t~~Three assessors carried out the assessments. Each trial was allocated to two assessors who independently applied the criteria. ~~team members who carried out the assessments are all NETSCC researchers with higher health degrees.~~ None of the ~~researchers~~ assessors had medical or clinical experience, however have higher health degrees and ~~,~~ work full time in health research and in evaluating clinical research. All NIHR HTA reports were examined by using a stabilised process; initially scanning the executive summary, followed by the methods, using key word search terms to scan the whole document and appendices, and finally undertaking a detailed reading of the entire ~~journal~~ report if relevant information could not be found. ~~Table 3 contains selected examples against six of the items in the checklist to illustrate both complete and poorly described interventions.~~

Data analysis

~~Disagreements in the scoring of reports were discussed in team meetings where differences in assessments were considered and resolved.~~

The checklist for each trial were completed using an electronic, stand-alone Access database. All three assessors completed checklists were then merged and exported into Excel and ~~Data were exported from Microsoft Access to~~ IBM SPSS version 19 for data analysis. IBM SPSS~~, and this~~ software was used to conduct all descriptive and inferential analyses. The chi-square test was used for all comparisons (statistically significant at $P < 0.05$). If any cell had an expected count less than 5, the Fisher's exact test was used.

Formatted: Font: Bold

Formatted: Plain Text, Justified

Results:

The modified checklist was applied to 98 RCTs published the HTA Journal series from January 1999 until March 2011. The interventions within each published trial were classified by the following intervention types; pharmaceutical, radiotherapy, surgery, diagnostic, education and training, service delivery, psychological, vaccines and biological, devices, physical therapy, exercise, complementary therapy, mixed or complex and other. The intervention classification was provided by *Schroter et al*[4] as part of the original checklist. Table ~~two~~2 shows the number of ~~journals~~trials within the journal series for each intervention by type.

Formatted: Font: Italic

Applying the ~~revised~~ modified checklist to NIHR HTA funded RCTs revealed that components of the intervention description were missing in ~~68~~ 68 of the~~/~~ 98 reports (missing 69.4%) ~~(Table 4)~~. Table 3 contains selected examples against six of the items in the checklist to illustrate both complete and poorly described interventions.

Intervention descriptions were therefore complete in 30.6% of reports. Certain criteria had high levels of completeness, such as baseline characteristics (94.9%) and descriptions of settings (91.8%), which were complete for over 90% of reports. However, other criteria were notably less complete, particularly patient information with only 58.2% having an adequate description (Table 4).

Differences in completion rates were noted between the fourteen types of interventions. For example, descriptions of interventions were more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively. The ~~, although cChi- Ssquare analysis~~ test showed that this difference was not statistically significant (p=0.77). Furthermore, this was not the case with certain criteria, such as baseline characteristics (drugs 93.3%, non-drugs 94.4%) and provider information (drugs 73.3, non-drugs 77.8%) where levels of completeness were higher in non-drug trials than drug interventions.

Descriptions of interventions were found to be least complete for psychological interventions with only 27.3% of RCTs in this area being complete. The chi~~, although Chi~~ squared ~~analysis~~ test revealed that this difference was not statistically significant when compared with drug interventions (p=1.00). Again, there were a few occasions where certain criteria had the highest levels of completeness of all intervention types, in particular with baseline characteristics and provider information with 100% and 90.9% of completeness respectively (Table 4).

The modified checklist included a question around the completeness of the control group. This was not a detailed evaluation of all the components of the control group but a broad assessment of whether the description appeared to be complete or not. ~~Data were collected on the completeness rates for Control Group information but~~ Given the interpretative nature of this question, control group information were not ~~included~~ included with the full data. ~~with the full data set as this criterion was not included in Schroter's original checklist.~~ The data revealed that 51% of RCTs had complete descriptions of control groups.

Discussion:

Statement of principle findings

This study has revealed that 30.6% (30/98) of studies with a single trial published in the HTA Journal have a full description of the intervention. The interventions described in published RCTs performed well against certain criteria, such as baseline characteristics (with 95% having an adequate description), but less well on other criteria, such as patient information (with 58% having an adequate description). Drug trials were slightly more complete than non-drug trials and psychological interventions with 33.3% of journals having a complete intervention description, although these differences were not statistically significant.

Strengths and weaknesses

The strengths of this study are that externally generated and tested criteria were applied to evaluate the effectiveness of intervention descriptions in NIHR HTA Programme funded RCTs. However, there were limitations. First, none of the assessors applying the criteria were medically trained, however, assessors were not commenting on the suitability of an intervention for use in practice but whether aspects of the description that would be required for use in practice were present. There is a possibility that someone with medical training would score the projects differently. In previous work the authors have been medically trained. Second, authors of the reports were not contacted to provide additional information beyond that provided in the publication. ~~As Glasziou has demonstrated,~~Previous studies have demonstrated that contacting the research teams or additional searches for intervention details does increase the completeness of intervention descriptions[3,9]. However it is questionable whether having to undertake additional searches outside the publication effectively enhances the ease of replicating study findings. A third limitation concerned the type of data being collected. Whilst all the criteria are dichotomous (in that they are all yes / no answers), the justification behind this categorisation has different degrees of interpretation. This could have resulted in overly-harsh assessments of completeness for certain criteria. For example, the recipient criterion is clear (are inclusion/exclusion criteria present) whilst ~~compared with the~~ greater interpretation is required ~~by~~ for the materials criterion which requires the assessor to determine if the description of the physical materials is adequate and therefore open to interpretation. Certainly the completion rate for materials was among the lowest across all studies with 58% and 69% completion rates for informational and physical materials respectively.

A further limitation of the study was that the checklist was not fully applied to~~a full assessment of~~ the control group of the published trials. T~~was not undertaken. The checklist could have been applied to the control group in addition to the intervention; t~~his would have

provided a more complete picture of how well controls are described within a study. Another limitation was that the number of journals assessed for completeness was very small for certain assessments (for example only 11 / 98 journals reported psychological interventions). It is possible therefore, that certain findings of completeness rate occurred by chance.

Meanings of the study

It is tempting to make comparisons with others studies assessing the usability of intervention descriptions. In particular, *Glasziou et al*[3] reported that 41/80 (51%) of published reports of single randomised trials and systematic reviews in popular journals were complete compared with 30/98 (30.6%) completeness of NIHR HTA funded RCT trials. Similarly, interventions in NIHR HTA reports appeared to be described less well than the 51 trials published in the BMJ assessed by *Schroter et al.* where 43% (22/51) of papers were considered to be of sufficient description to allow replication[4]. Whilst these comparisons are interesting, it is important to note that it is not possible to make any meaningful comparison on the relative performance of each output, as the *Glasziou et al*[3] study looked at journal articles and we looked at the HTA journal series which are aimed at different audiences and the questionnaire used was different between the studies. This is because the nature of outputs varies considerably between studies as does the assessment criteria. It is notable, for example, that *Schroter et al*[4] used ~~seven~~ eight indicators (seven main checklist items and a global completeness eighth item) in their checklist, compared with the twelve criteria used in this study.

However, this study does reflect findings from similar studies conducted elsewhere. For example, the criteria highlighted as being particularly poorly described in Schroter's study were physical / informational materials, which reflected findings in this study where patient information and physical materials were also lacking in completeness. Similarly the fact that NIHR HTA Programme funded drug interventions were typically better described than non-drug interventions reflected findings in *Glasziou et al*[3]. where over 60% of reports on drug treatments were initially deemed to be complete compared with just under 30% of non- drug treatments.

In addition to the more detailed guidance provided to authors, the HTA Journal requests that authors of RCTs include the headings set out in the revised CONSORT checklist and flowchart and provide details of CONSORT in its guidance for authors. Item five of the CONSORT statement says 'The interventions for each group with sufficient details to allow replication, including how and when they were actually administered' and there are extensions of the CONSORT statement to address the additional complexity around the

reporting of non-pharmacological interventions. The CONSORT extensions are not currently a requirement for non-pharmacological studies but as these extensions are more widely requested, it is hopeful that the reporting of interventions will improve and be fully described.

Understanding the extent to which interventions in published studies are described sufficiently to inform clinical decision making is a key concern in the adding value in research agenda. As Chalmers and Glasziou have suggested, poorly described interventions form one of the four main pillars of research waste[1]. The criteria identified by *Schroter et al*[4] and developed in this study are helpful in highlighting specific areas of where intervention descriptions can be improved.

Future research

Several areas for further research are indicated by this study. Further testing on the criteria can be undertaken to assess the repeatability of the criteria. For example, the reports sampled in this study could be reassessed by someone with clinical experience to assess the level of agreement. Alternatively, Glasziou's selected papers in his original study could be assessed by non-clinical teams to examine the level of agreement. The checklist has only been applied to single trial studies; future research into the applicability of it for multi-trial studies should be investigated.

The characterisation of the control group is a key area for future research, as research involving trials to date has focused on the description of interventions with a treatment group, however the detail of the control arm is equally important as in many cases the control arm is often described as 'usual care' but this does not take into account variations by centre[10]. A recent paper reports[11] a development of a tool for extraction of data in systematic reviews and includes an element on intervention design. The tool has been applied to both the intervention and control groups of systematic reviews and the applicability of its use across primary research should be investigated.

Ensuring the replicability of study findings is an essential part of adding value in research. It is important for health research publishers to be transparent in the usability of study reports and areas of improvement. This study applied a checklist that can be used to indicate where the descriptions of interventions can be improved to enhance replication in clinical practice. Serious consideration should be given on how this might be used to improve intervention reporting in the future. The results of this study have been shared with the editorial Board of the HTA Journal to investigate how interventions can be better reported within the journal series.

References

1. Chalmers I,Glasziou P. Avoidable waste in the production and reporting of research evidence. *The Lancet* 2009;**374**:86-9
2. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ* 2010; 341.
3. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008; 336(7659):1472-1474.
4. Schroter S, Glasziou P, Heneghan C. Quality of description of treatments: a review of published randomised controlled trials. BMJ Open 2012; 2.
5. Thabane L, Chu R, Cuddy K, Douketis J. What is the quality of reporting in weight loss intervention studies? A systematic review of randomized controlled trials. [Review] [17 refs]. *International Journal of Obesity* 2007; 31(10):1554-1559
6. Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in Hodgkin's lymphoma in biomedial journals. *J Natl Cancer Inst* 2006; 98:620-625.
7. Lai R, Chu R, Fraumeni M, et al. Quality of ransomized controlled trials reporting in the primary treamt of brain tumors. *J Clin Oncol* 2006; 24:1136-1144.
8. Glenton C, Underland V, Kho M, Pennick V, Oxman AD. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. *Journal of Clinical Epidemiology* 2006; 59(8):770-778
9. Hoffmann T, Erueti C, Glasziou P. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ* 2013; 347:f3755
10. Cook A, Douet L and Boutron I. Descriptions of non-pharmacological interventions in clinical trials BMJ 2013;347:f5212
11. Montgomerya P, Underhill K, Gardnera F, Operarioc D, Mayo-Wilson E. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology* 2013 66(8): 874-882

Table 1: Replicability Criteria for Interventions, developed from an initial design by Schroter et al[3]

Checklist criteria		Descriptor of criteria where appropriate
1. Setting	Is it clear where the intervention was delivered?	
2a. Recipient – inclusion	Is it clear who is receiving the intervention? – Inclusion criteria	Clear inclusion criteria in the journal
2b. Recipient – exclusion	Is it clear who is receiving the intervention? – Exclusion criteria	Clear exclusion criteria in journal
2c. Recipient – baseline characteristics	Do you know all that you need to about the patients? (e.g. which drugs they are taking, what they were told, etc)? If No, what further information do you require?	Baseline characteristics of participants provided in journal
3. Provider	Is it clear who delivered the intervention?	
4. Procedure	Is the procedure (including the sequencing of the technique) of the intervention sufficiently clear to allow replication?	Top level overview of the intervention. Eg Drug X for X days at X dose. Or X sessions lasting X minutes, for X weeks/ months
5. Intensity	Is the dose/duration of individual sessions of the intervention clear?	Dose, length of session
6. Schedule	Is the schedule (interval, frequency, duration, or timing) of the intervention clear?	Frequency of intervention, length of session
7a. Materials – physical	Are the physical materials used adequately described?	Physical materials eg Description of splint used. If either are no, it is a no overall
7b. Materials - informational	Are the informational materials used adequately described?	Information provided to the patients eg consent forms etc
8. Missing	Is the description of the intervention complete? If No, what is missing?	
9. Control	Is it clear what the control group received during the study?	

Table 2: Intervention type of NIHR HTA funded trial RCTs included in the study

Type of intervention	№ (%)
Drug	15 (15.3)
Radiotherapy	1 (1.0)
Surgery	9 (9.2)
Diagnostic	8 (8.2)
Education and training	3 (3.1)
Service delivery	19 (19.4)
Psychological therapies	11 (11.2)
Vaccines and biologicals	3 (3.1)
Devices	12 (12.2)
Physical therapies	7 (7.1)
Exercise	1 (1.0)
Complementary therapies	2 (2.0)
Mixed or complex	6 (6.1)
Other*	1 (1.0)
Total	**98**

*Other refers to an intervention using larval therapy

Table 3: Examples of poor reporting of intervention elements within the HTA journal series, taken verbatim form the journal.

Checklist item	Examples of poor reporting	Reason why rated as incomplete	Examples of good reporting	Reason why rated complete
Inclusion criteria	'….patient identification was retrospective. Searches were conducted on practice databases using either repeat prescriptions alone or repeat prescriptions plus diagnostic terms… GPs then sent letters to suitable patients, providing information about the trial'	No details given about the searches and the criteria patients were screened with.	'Inclusion criteria for trial patients were: -Diagnosed with idiopathic arthritides of childhood with onset before their 16th birthday for more than 3 months. - Aged 4–19 years inclusive. - Stable on medication. - At least one active joint, core set criteria 1.56 -At least two out of any five of the remaining core set criteria below. -The physician global assessment of disease activity >10 mm on a 100-mm visual analogue scale (VAS). - The parent global assessment of well-being >10 mm on a 100-mm VAS. -Childhood Health Assessment Questionnaire scores >0. - More than one joint with limited range of motion (joint motion reduced by at least 5° from normative range for age58). -An elevated erythrocyte sedimentation rate (ESR) (>5 mmHg in children and >10 mmHg in adolescents)'	Very detailed patient criteria listed
Exclusion criteria	'GPs were given a ringbinder file with information and instructions about the trial and, within each, a number of recruitment packs. The packs contained the paperwork required to complete the recruitment of each patient, this was: a reminder of the inclusion/exclusion criteria for the study….'	No details given about the exclusion criteria	'Reasons for Exclusion (Yes/No) - BMI > 40 kg/m2 - Barrett's oesophagus (≥3cm) - Paraoesophageal hernia - Oesophageal strictures - One type of management is clinically indicated for another reason'	Detailed patient exclusion criteria listed
Provider	'All services had staff who were trained and experienced in family therapy, but not necessarily family interventions specifically for eating disorders'	No details about the staff providing the interventions or the training they received.	'Eight counsellors (six females and two males) took part in the trial (one worked at two practices) and all were BAC accredited or eligible for BAC accreditation; they were highly trained and had considerable experience of counselling in a general practice setting' (There are details about each counsellors age, qualifications and experience are provided)	States who delivered the intervention and their training
Procedure	'Generally home-based rehabilitation services provide, as a minimum, physiotherapy and occupational therapy in the patient's own home. Services can be specialised (e.g. in stroke rehabilitation) or be provided for patients with a range of disabilities'	No details about the services provided to patients and variation between centres	'The content of the CBT programme included (complete course description contained within an appendix): - Elucidation of core beliefs regarding their illness and its management. - Monitoring of activity levels and introduction of appropriate timetable. - Introduction to exercises designed to increase general level of fitness, balance and confidence in exercise. A range of aerobic, strength, balance and stretching exercises were taught. - Behavioural modification of sleep patterns. - Mood management advice. - Goal setting'	Key aspects of the intervention summarised in the text and a full description of the intervention is detailed in the appendices.
Intensity and Schedule	'Patients come to the day hospital where the rehabilitation service is provided for a full or half day. Usually ambulance transport is provided to bring patients into the service and return them home after a session'	No details of the length or number of sessions	'Psychological treatment was based on existing protocols (references included) and distributed over six 50-minute sessions, with printed information sheets provided after each session'	The length and number of sessions is included as well as the details of each session.
Materials – physical	'The acupuncture point prescriptions used were individualised to each patient and were at the discretion of the acupuncturist'	The prescriptions used are not detailed.	'- 500 mg oral oxytetracycline (non-proprietary) b.d. + topical vehicle control b.d. -100 mg oral Minocin MR minocycline) o.d. + topical vehicle control b.d. -topical Panoxyl Aquagel (5% benzoyl peroxide) b.d. + oral placebo o.d. This was designated as the active comparator group, as benzoyl peroxide was the leading and most established topical treatment for acne when the protocol was written. -topical Benzamycin (3% erythromycin + 5% benzoyl peroxide) b.d. + oral placebo o.d. (referred to as ery. + BP bd) -topical Stiemycin (2% erythromycin) o.d. +topical Panoxyl Aquagel (5% benzoyl peroxide)o.d. + oral placebo o.d. (referred to as ery. od+ BP od)'	Each of the treatments prescribed is clearly defined

Table 4: Completion rates of NIHR HTA reports by intervention type

Description criteria	Drugs No. Trials with complete intervention description (%)	CI*	Non-Drugs No. Trials with complete intervention description (%)	CI*	Psychological No. Trials with complete intervention description (%)	CI*	All No. Trials with complete intervention description (%)
Number journals	15 (15.3)		72 (73.5)		11 (11.2)		98 (100)
Setting	15 (100)	0.80-1	66 (91.7)	0.83-0.96	9 (81.8)	0.52-0.95	90 (91.8)
Inclusion criteria	15 (100)	0.80-1	63 (87.5)	0.78-0.93	10 (90.9)	0.62-0.98	88 (89.8)
Exclusion criteria	14 (93.3)	0.70-0.99	55 (76.4)	0.65-0.85	10 (90.9)	0.62-0.98	79 (80.6)
Baseline characteristics	14 (93.3)	0.70-0.99	68 (94.4)	0.87-0.98	11 (100)	0.74-1	93 (94.9)
Provider	11 (73.3)	0.48-0.89	56 (77.8)	0.67-0.86	10 (90.9)	0.62-0.98	77 (78.6)
Procedure	14 (93.3)	0.70-0.99	57 (79.2)	0.68-0.87	9 (81.8)	0.52-0.95	80 (81.6)
Intensity	13 (86.7)	0.62-0.96	63 (87.5)	0.78-0.93	9 (81.8)	0.52-0.95	85 (86.8)
Schedule	13 (86.7)	0.62-0.96	59 (81.9)	0.71-0.89	9 (81.8)	0.52-0.95	81 (82.7)
Patient information	9 (60.0)	0.36-0.80	42 (58.3)	0.47-0.69	6 (54.5)	0.28-0.79	57 (58.2)
Physical materials	10 (66.7)	0.42-0.85	52 (72.3)	0.61-0.81	6 (54.5)	0.28-0.79	68 (69.4)
Intervention description complete overall	5 (33.3)	0.15-0.58	22 (30.6)	0.21-0.42	3 (27.3)	0.10-0.57	30 (30.6)

**Please note: some criteria are not applicable therefore denominator less than total number of journals.*95% confidence interval: no continuity correction*

The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-003713.R2
Article Type:	Research
Date Submitted by the Author:	28-Nov-2013
Complete List of Authors:	Douet, Lisa; University of Southampton, National Institute for Health Research Milne, Ruairidh; University of Southampton, Wessex Institute Anstee, Sydney; University of Southampton, National Institute for Health Research Habens, Fay; University of Southampton, National Institute for Health Research Young, Amanda; University of Southampton, NETSCC Wright, David; University of Southampton, National Institute for Health Research
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Research methods
Keywords:	Clinical trials < THERAPEUTICS, AUDIT, STATISTICS & RESEARCH METHODS

The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Dr Lisa Douet[a]
Senior Research Fellow

Professor Ruairidh Milne[b]
Professorial Fellow in Public Health

Sydney Anstee[a]
Research Fellow

Fay Habens[a]
Programme Manager

Amanda Young[a]
Senior Research Fellow

Dr David Wright[a]
Senior Research Fellow

[a]National Institute for Health Research
Evaluation, Trials and Studies Coordinating Centre (NETSCC)
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

[b]Wessex Institute
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

Correspondence to: l.douet@southampton.ac.

Abstract word count: 281

Main text word count: 3105

Abstract:

Objectives

The objective of this study is to assess whether NIHR HTA funded randomised controlled trials (RCTs) published in *Health Technology Assessment* Journal were described in sufficient detail to replicate in practice.

Methods

A published checklist for assessing intervention descriptions was applied to NIHR HTA funded RCTs published in *Health Technology Assessment*. The checklist was piloted twice on a sample of 10 reports and modified. Kappa scores were generated to assess agreement in the checklist application. The modified checklist was applied to all 98 NIHR HTA funded single trial RCTs published in the Journal up to March 2011. The checklist included assessments of participants, intensity, schedule, materials and settings. A study was agreed to be complete overall, is aspects of the checklist were present. Assessors independently applied the checklist. Disagreements in scoring were discussed in the team; differences were then explored and resolved.

Results

Components of the intervention description were missing in 68 / 98 (69.4%) reports. Baseline characteristics and descriptions of settings had the highest levels of completeness with over 90% of reports complete. Reports were less complete on patient information with 58.2% of the journals having an adequate description. When looking at individual intervention types, drug intervention descriptions were more complete than non-drug interventions with 33.3% and 30.6% levels of completeness respectively, although this was not significant statistically. Only 27.3% of RCTs with psychological interventions were deemed to be complete, although again these differences were not significant statistically.

Conclusions

Ensuring the replicability of study interventions is an essential part of adding value in research. All those publishing clinical trial data need to ensure transparency and completeness in the reporting of interventions to ensure that study interventions can be replicated.

Article summary

Article focus

- This article investigates the adequacy of reporting of different elements of interventions reported in randomised controlled trials published within the NIHR HTA journal series, using a simple checklist

Key messages

- Only 30.6% (30/98) of studies with a single trial published in the NIHR HTA Journal have a full description of the intervention
- The unlimited word count of the NIHR journal series does not affect completeness of an intervention description

Strengths and limitations

- An externally produced checklist was applied to all RCTs publishing in the NIHR HTA Journal Series
- The sample size for a number of assessments is very small

Introduction:

A recent publication by Glasziou and Chalmers has suggested that as much as 85% of the US$100 billion spent on health research worldwide each year is potentially wasted due to four key problems of knowledge production and dissemination[1].

Several studies have specifically assessed the waste area of ensuring funded research is unbiased and usable by exploring the quality and usability of publications from funded health research. This is a key concern considering the role effective summaries of evidence have in facilitating knowledge transfer and enhancing the uptake of findings in clinical practice. Whilst it is recognised that trial registration databases and scientific journals can be restrictive in terms of word allowance, various strategies have been proposed to improve the reporting of interventions in published trials, including an 'intervention bank' to include manuals and fidelity tools linked to trial registration numbers'[2].

Studies have highlighted concerns about the descriptions of interventions in final reports and publications. In one study, for example, 80 consecutive studies were selected for assessment of completeness from the journal Evidence-Based Medicine. Two general practitioners independently assessed whether they could use the treatment with a patient if they saw them tomorrow[3]. Of these 80 published reports, 41 (51%) of had elements of the intervention missing, particularly descriptions of process and information on hand-outs or booklets. The proportion of trials for which adequate information could be made available increased to 90% through the checking of references, contacting authors, and undertaking additional searches[3].

Similarly, *Schroter et al* developed, piloted and applied a checklist designed to assess the replicability of published treatment decisions to 51 trials published in the BMJ[4]. This checklist was applied by the study team to a broad range of health topics and included seven items and a global eighth item to summarise completeness. This study reported that 57% (29/51) of papers were not considered to be of sufficient description to allow replication, with the most poorly described aspects of the published trials being the sequencing of the technique and physical / information materials[4]. A further study[9] has used the checklist developed by *Schroter et al* to assess the completeness of non-pharmacological intervention description and reported that only 39% were adequately described.

Rates of replicability of interventions vary considerably in the published literature depending on the complexity of the treatment and the assessment criteria. For example, three studies assessed compliance with item four of CONSORT in published research in the areas of

weight loss[5], brain tumours[6] and Hodgkin's lymphoma[7]. Item four of CONSORT specifically asks for precise details concerning treatments intended for all groups and how and when they were administered. These studies reported that over 90% of study findings were replicable. In contrast, however, one study assessed whether there was sufficient information on what happens before, during and after treatment for back pain, and revealed that only 13% of trials were replicable[8].

The NIHR Health Technology Assessment (HTA) Programme commissions and funds primary research and evidence synthesis on the effectiveness, costs and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the NHS. It aspires to enable all funded projects to complete and publish in the programme's own Journal '*Health Technology Assessment*', freely available on the programme's website (www.hta.ac.uk). Reports published in the Journal series are peer reviewed, are in the public domain and contain a full record of the study. Unlike typical peer reviewed journals, there are no word or size limitations for the full report and unlimited appendices, thus enabling more detail to be included in the publication; an average report is approximately 50,000 in length. Given the importance of complete and replicable reporting of findings and the opportunities the NIHR HTA Journal presents, this study aimed to assess whether randomised controlled trials (RCTs) with single trials published in *Health Technology Assessment* were described in sufficient detail.

Methods:

Data source

All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion in the study. Of the 109 reports publishing in this time-period, 11 were excluded as they reported more than one RCT within a single HTA Journal. Ninety-eight single trial RCTs were therefore included in the study.

Piloting the checklist

Five NIHR HTA funded RCTs were selected to represent a range of interventions (surgery, psychology, devices and pharmaceutical) to pilot the checklist initially developed by *Schroter et al*[4]. The checklist was applied independently by three assessors to assess the level of agreement between assessments. A kappa score was produced for each individual trial (range 0.15 – 0.7) with an average of 0.225 across all 5 trials. Disagreement was due to differing interpretations of the checklist questions. The initial checklist was modified to separate out two of the questions into their individual components. In the published checklist the recipient question stated 'Is it clear who is receiving the intervention?' and 'Do you know all that you need to about the patients? (e.g. which drugs they are taking, what they were told, etc)?'. We felt that this required several pieces of information for a single question and therefore we separated question 2 into the three components, as shown in table 1. Similarly the materials question, 'Are the physical or informational materials used adequately described?' was separated into the two components shown in question 7 of table 1. Additionally, the assessors discussed the type and level of information expected to be present in order to answer a question as complete. The modified checklist was applied by the three assessors to a further five NIHR HTA funded RCTs, resulting in higher levels of agreement (kappa scores for each report ranged from 0.3 – 0.7, with an average of 0.6 for all trials).

The main study

The final modified checklist was applied to a wider sample of NIHR HTA funded RCTs. All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion with in the study. One checklist was completed for the intervention group of each trial published. Each item in the checklist was answered by either a yes, no or not applicable response. We did not apply the full checklist to the Control Group, but, unlike the published checklist (Schroter et al[4]), we did make a general assessment as the to completeness of control group information within question nine of the checklist. However, responses to this question were not on a detailed assessment of all components of the

control group, unlike the intervention group itself. Question eight summarises whether there are any aspects of the intervention missing based on the responses to the previous seven questions.

Data quality

Each trial was assessed independently by two assessors. 15% of published reports (15/98) were discussed due to disagreements of the scoring mainly around checklist item seven. All disagreements were discussed by the team and were resolved by consensus.

Three assessors carried out the assessments. Each trial was allocated to two assessors who independently applied the criteria. None of the assessors had medical or clinical experience; however have higher health degrees and work full time in health research and in evaluating clinical research. All NIHR HTA reports were examined by using a stabilised process; initially scanning the executive summary, followed by the methods, using key word search terms to scan the whole document and appendices, and finally undertaking a detailed reading of the entire report if relevant information could not be found.

Data Analysis

The checklist for each trial was completed using an electronic, stand-alone Access database. All three assessors completed checklists were then merged and exported into Excel and IBM SPSS version 19 for data analysis. IBM SPSS software was used to conduct all descriptive and inferential analyses. The chi-square test was used for all comparisons (statistically significant at P <0.05). If any cell had an expected count less than 5, the Fisher's exact test was used.

Results:

The modified checklist was applied to 98 RCTs published the HTA Journal series from January 1999 until March 2011. The interventions within each published trial were classified by the following intervention types; pharmaceutical, radiotherapy, surgery, diagnostic, education and training, service delivery, psychological, vaccines and biological, devices, physical therapy, exercise, complementary therapy, mixed or complex and other. The intervention classification was provided by *Schroter et al*[4] as part of the original checklist. Table 2 shows the number of trials within the journal series for each intervention by type.

Applying the modified checklist to NIHR HTA funded RCTs revealed that components of the intervention description were missing in 68 of the 98 reports (missing 69.4%). Table 3 contains selected examples against six of the items in the checklist to illustrate both complete and poorly described interventions.

Intervention descriptions were therefore complete in 30.6% of reports. Certain criteria had high levels of completeness, such as baseline characteristics (94.9%) and descriptions of settings (91.8%), which were complete for over 90% of reports. However, other criteria were notably less complete, particularly patient information with only 58.2% having an adequate description (Table 4).

Differences in completion rates were noted between the fourteen types of interventions. For example, descriptions of interventions were more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively. The chi-square test showed that this difference was not statistically significant (p=0.77). Furthermore, this was not the case with certain criteria, such as baseline characteristics (drugs 93.3%, non-drugs 94.4%) and provider information (drugs 73.3, non-drugs 77.8%) where levels of completeness were higher in non-drug trials than drug interventions.

Descriptions of interventions were found to be least complete for psychological interventions with only 27.3% of RCTs in this area being complete. The chi-squared test revealed that this difference was not statistically significant when compared with drug interventions (p=1.00). Again, there were a few occasions where certain criteria had the highest levels of completeness of all intervention types, in particular with baseline characteristics and provider information with 100% and 90.9% of completeness respectively (Table 4).

The modified checklist included a question around the completeness of the control group. This was not a detailed evaluation of all the components of the control group but a broad assessment of whether the description appeared to be complete or not. Given the

interpretative nature of this question, control group information were not included with the full data. The data revealed that 51% of RCTs had complete descriptions of control groups.

Discussion:

Statement of principle findings

This study has revealed that 30.6% (30/98) of studies with a single trial published in the HTA Journal have a full description of the intervention. The interventions described in published RCTs performed well against certain criteria, such as baseline characteristics (with 95% having an adequate description), but less well on other criteria, such as patient information (with 58% having an adequate description). Drug trials were slightly more complete than non-drug trials and psychological interventions with 33.3% of journals having a complete intervention description, although these differences were not statistically significant.

Strengths and weaknesses

The strengths of this study are that externally generated and tested criteria were applied to evaluate the effectiveness of intervention descriptions in NIHR HTA Programme funded RCTs. However, there were limitations. First, none of the assessors applying the criteria were medically trained, however, assessors were not commenting on the suitability of an intervention for use in practice but whether aspects of the description that would be required for use in practice were present. There is a possibility that someone with medical training would score the projects differently. In previous work the authors have been medically trained. Second, authors of the reports were not contacted to provide additional information beyond that provided in the publication. Previous studies have demonstrated that contacting the research teams or additional searches for intervention details does increase the completeness of intervention descriptions[3,9]. However it is questionable whether having to undertake additional searches outside the publication effectively enhances the ease of replicating study findings.

A limitation of the checklist used is the type of data being collected. Whilst all the criteria are dichotomous (in that they are all yes / no answers), the justification behind this categorisation has different degrees of interpretation. This could have resulted in overly-harsh assessments of completeness for certain criteria. For example, the recipient criterion is clear (are inclusion/exclusion criteria present) whilst greater interpretation is required for the materials criterion which requires the assessor to determine if the description of the physical materials is adequate and therefore open to interpretation. Certainly the completion rate for materials was among the lowest across all studies with 58% and 69% completion rates for informational and physical materials respectively. By using this checklist we were able to suggest further refinements to the criteria used within it, such as separating out the recipient criteria and the material criteria.

A further limitation of the study was that the checklist was not fully applied to the control group of the published trials. This would have provided a more complete picture of how well controls are described within a study. Another limitation was that the number of journals assessed for completeness was very small for certain assessments (for example only 11 / 98 journals reported psychological interventions). It is possible therefore, that certain findings of completeness rate occurred by chance.

Meanings of the study

It is tempting to make comparisons with others studies assessing the usability of intervention descriptions. In particular, *Glasziou et al*[3] reported that 41/80 (51%) of published reports of single randomised trials and systematic reviews in popular journals were complete compared with 30/98 (30.6%) completeness of NIHR HTA funded RCT trials. Similarly, interventions in NIHR HTA reports appeared to be described less well than the 51 trials published in the BMJ assessed by *Schroter et al.* where 43% (22/51) of papers were considered to be of sufficient description to allow replication[4]. Whilst these comparisons are interesting, it is important to note that it is not possible to make any meaningful comparison on the relative performance of each output, as the *Glasziou et al*[3] study looked at journal articles and we looked at the HTA journal series which are aimed at different audiences and the questionnaire used was different between the studies. This is because the nature of outputs varies considerably between studies as does the assessment criteria. It is notable, for example, that *Schroter et al*[4] used eight indicators (seven main checklist items and a global completeness eighth item) in their checklist, compared with the twelve criteria used in this study.

However, this study does reflect findings from similar studies conducted elsewhere. For example, the criteria highlighted as being particularly poorly described in Schroter's study were physical / informational materials, which reflected findings in this study where patient information and physical materials were also lacking in completeness. Similarly the fact that NIHR HTA Programme funded drug interventions were typically better described than non-drug interventions reflected findings in *Glasziou et al*[3]. where over 60% of reports on drug treatments were initially deemed to be complete compared with just under 30% of non- drug treatments.

In addition to the more detailed guidance provided to authors, the HTA Journal requests that authors of RCTs include the headings set out in the revised CONSORT checklist and flowchart and provide details of CONSORT in its guidance for authors. Item five of the CONSORT statement says 'The interventions for each group with sufficient details to allow

replication, including how and when they were actually administered' and there are extensions of the CONSORT statement to address the additional complexity around the reporting of non-pharmacological interventions. The CONSORT extensions are not currently a requirement for non-pharmacological studies but as these extensions are more widely requested, it is hopeful that the reporting of interventions will improve and be fully described.

A number of studies have investigated the completeness of intervention descriptions in a single disease area by assessing compliance of RCTs with the intervention item (item 4) of the CONSORT statement[5,6,7]. Whilst these studies reported that over 90% of study findings were replicable, it is likely that this is an over estimation as they do not assess the question of whether there was enough information to allow replication. In contrast, however, one study assessed whether there was sufficient information on what happens before, during and after treatment for back pain, and revealed that only 13% of trials were replicable[8].

Understanding the extent to which interventions in published studies are described sufficiently to inform clinical decision making is a key concern in the adding value in research agenda. As Chalmers and Glasziou have suggested, poorly described interventions form one of the four main pillars of research waste[1]. The criteria identified by *Schroter et al*[4] and developed in this study are helpful in highlighting specific areas of where intervention descriptions can be improved.

Future research

Several areas for further research are indicated by this study. Further testing on the criteria can be undertaken to assess the repeatability of the criteria. For example, the reports sampled in this study could be reassessed by someone with clinical experience to assess the level of agreement. Alternatively, Glasziou's selected papers in his original study could be assessed by non-clinical teams to examine the level of agreement. The checklist has only been applied to single trial studies; future research into the applicability of it for multi-trial studies should be investigated.

The characterisation of the control group is a key area for future research, as research involving trials to date has focused on the description of interventions with a treatment group, however the detail of the control arm is equally important as in many cases the control arm is often described as 'usual care' but this does not take into account variations by centre[10]. A recent paper reported on the development of a tool for extraction of data in systematic reviews and includes an element on intervention design[11]. The tool has been applied to both the intervention and control groups of systematic reviews. The applicability of the tool

across primary research could be investigated and used to further strengthen the checklist that we have used.

Ensuring the replicability of study findings is an essential part of adding value in research. It is important for health research publishers to be transparent in the usability of study reports and areas of improvement. This study applied a checklist that can be used to indicate where the descriptions of interventions can be improved to enhance replication in clinical practice. Serious consideration should be given on how this might be used to improve intervention reporting in the future. The results of this study have been shared with the editorial Board of the HTA Journal to investigate how interventions can be better reported within the journal series.

Acknowledgements:
We would like to acknowledge Professor James Raftery and the Metadata team for providing the database and the trial details used in the study and for Professor Paul Glasziou for his advice during the study.

Ethic statement:
This study did not require ethics approval as no data patient or clinical data was required.

Funding statement:
This study was supported by the NIHR Evaluation, Trials and Studies Coordinating Centre through its Research on Research Programme. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Department of Health, or of NETSCC.

Contribution of Authors:
The study was conceived and designed by LD, RM, SA, and FH, and undertaken by LD, SA and FH; AY and DW supported the data analysis. All authors read and approved the final manuscript. LD is guarantor of the study

Data sharing: Data on the included trials are available on request from the corresponding author.

Competing Interests: The authors have no competing financial interests; however, all of the authors are employed by the University of Southampton to work at least part time for NETSCC. In particular: RM is employed as the Head of NETSCC and has worked for NETSCC (and its predecessor organisation) in senior roles on and off since 1996. He was an editor of the Health Technology Assessment journal (1997-2007) and a founder editor for other journals in the new NIHR Journals Library (2011-12).

References

1. Chalmers I,Glasziou P. Avoidable waste in the production and reporting of research evidence. *The Lancet* 2009;**374**:86-9
2. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ* 2010; 341.
3. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008; 336(7659):1472-1474.
4. Schroter S, Glasziou P, Heneghan C. Quality of description of treatments: a review of published randomised controlled trials. BMJ Open 2012; 2.
5. Thabane L, Chu R, Cuddy K, Douketis J. What is the quality of reporting in weight loss intervention studies? A systematic review of randomized controlled trials. [Review] [17 refs]. *International Journal of Obesity* 2007; 31(10):1554-1559
6. Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in Hodgkin's lymphoma in biomedial journals. *J Natl Cancer Inst* 2006; 98:620-625.
7. Lai R, Chu R, Fraumeni M, et al. Quality of ransomized controlled trials reporting in the primary treamt of brain tumors. *J Clin Oncol* 2006; 24:1136-1144.
8. Glenton C, Underland V, Kho M, Pennick V, Oxman AD. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. *Journal of Clinical Epidemiology* 2006; 59(8):770-778
9. Hoffmann T, Erueti C, Glasziou P. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ* 2013; 347:f3755
10. Cook A, Douet L and Boutron I. Descriptions of non-pharmacological interventions in clinical trials BMJ 2013;347:f5212
11. Montgomerya P, Underhill K, Gardnera F, Operarioc D, Mayo-Wilson E. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology* 2013 66(8): 874-882

The completeness of intervention descriptions in published NIHR HTA funded trials: A cross sectional study

Dr Lisa Douet[a]
Senior Research Fellow

Professor Ruairidh Milne[b]
Professorial Fellow in Public Health

Sydney Anstee[a]
Research Fellow

Fay Habens[a]
Programme Manager

Amanda Young[a]
Senior Research Fellow

Dr David Wright[a]
Senior Research Fellow

[a]National Institute for Health Research
Evaluation, Trials and Studies Coordinating Centre (NETSCC)
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

[b]Wessex Institute
University of Southampton
Alpha House, Enterprise Road
Southampton SO16 7NS

Correspondence to: l.douet@southampton.ac.uk

Acknowledgements:
We would like to acknowledge Professor James Raftery and the Metadata team for providing the database and the trial details used in the study and for Professor Paul Glasziou for his advice during the study.

Ethic statement:
This study did not require ethics approval as no data patient or clinical data was required.

Funding statement:
This study was supported by the NIHR Evaluation, Trials and Studies Coordinating Centre through its Research on Research Programme. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Department of Health, or of NETSCC.

Contribution of Authors:
The study was conceived and designed by LD, RM, SA, and FH, and undertaken by LD, SA and FH; AY and DW supported the data analysis. All authors read and approved the final manuscript. LD is guarantor of the study

Data sharing: Data on the included trials are available on request from the corresponding author.

Competing Interests: The authors have no competing financial interests; however, all of the authors are employed by the University of Southampton to work at least part time for NETSCC. In particular: RM is employed as the Head of NETSCC and has worked for NETSCC (and its predecessor organisation) in senior roles on and off since 1996. He was an editor of the Health Technology Assessment journal (1997-2007) and a founder editor for other journals in the new NIHR Journals Library (2011-12).

Abstract word count: 281
Main text word count: ~~2992~~3105

Abstract:

Objectives

The objective of this study is to assess whether NIHR HTA funded randomised controlled trials (RCTs) published in *Health Technology Assessment* Journal were described in sufficient detail to replicate in practice.

Methods

A published checklist for assessing intervention descriptions was applied to NIHR HTA funded RCTs published in *Health Technology Assessment*. The checklist was piloted twice on a sample of 10 reports and modified. Kappa scores were generated to assess agreement in the checklist application. The modified checklist was applied to all 98 NIHR HTA funded single trial RCTs published in the Journal up to March 2011. The checklist included assessments of participants, intensity, schedule, materials and settings. A study was agreed to be complete overall, is aspects of the checklist were present. Assessors independently applied the checklist. Disagreements in scoring were discussed in the team; differences were then explored and resolved.

Results

Components of the intervention description were missing in 68 / 98 (69.4%) reports. Baseline characteristics and descriptions of settings had the highest levels of completeness with over 90% of reports complete. Reports were less complete on patient information with 58.2% of the journals having an adequate description. When looking at individual intervention types, drug intervention descriptions were more complete than non-drug interventions with 33.3% and 30.6% levels of completeness respectively, although this was not significant statistically. Only 27.3% of RCTs with psychological interventions were deemed to be complete, although again these differences were not significant statistically.

Conclusions

Ensuring the replicability of study interventions is an essential part of adding value in research. All those publishing clinical trial data need to ensure transparency and completeness in the reporting of interventions to ensure that study interventions can be replicated.

Article summary

Article focus

- This article investigates the adequacy of reporting of different elements of interventions reported in randomised controlled trials published within the NIHR HTA journal series, using a simple checklist

Key messages

- Only 30.6% (30/98) of studies with a single trial published in the NIHR HTA Journal have a full description of the intervention
- The unlimited word count of the NIHR journal series does not affect completeness of an intervention description

Strengths and limitations

- An externally produced checklist was applied to all RCTs publishing in the NIHR HTA Journal Series
- The sample size for a number of assessments is very small

Introduction:

A recent publication by Glasziou and Chalmers has suggested that as much as 85% of the US$100 billion spent on health research worldwide each year is potentially wasted due to four key problems of knowledge production and dissemination[1].

Several studies have specifically assessed the waste area of ensuring funded research is unbiased and usable by exploring the quality and usability of publications from funded health research. This is a key concern considering the role effective summaries of evidence have in facilitating knowledge transfer and enhancing the uptake of findings in clinical practice. Whilst it is recognised that trial registration databases and scientific journals can be restrictive in terms of word allowance, various strategies have been proposed to improve the reporting of interventions in published trials, including an 'intervention bank' to include manuals and fidelity tools linked to trial registration numbers'[2].

Studies have highlighted concerns about the descriptions of interventions in final reports and publications. In one study, for example, 80 consecutive studies were selected for assessment of completeness from the journal Evidence-Based Medicine~~, a journal aimed specifically at doctors working in primary care and general medicine~~. Two general practitioners independently assessed whether they could use the treatment with a patient if they saw them tomorrow[3]. Of these 80 published reports, 41 (51%) of had elements of the intervention missing, particularly descriptions of process and information on hand-outs or booklets. The proportion of trials for which adequate information could be made available increased to 90% through the checking of references, contacting authors, and undertaking additional searches[3].

Similarly, *Schroter et al* developed, piloted and applied a checklist designed to assess the replicability of published treatment decisions to 51 trials published in the BMJ[4]. This checklist was applied by the study team to a broad range of health topics and included seven items and a global eighth item to summarise completeness. This study reported that 57% (29/51) of papers were not considered to be of sufficient description to allow replication, with the most poorly described aspects of the published trials being the sequencing of the technique and physical / information materials[4]. A further study[9] has used the checklist developed by *Schroter et al* to assess the completeness of non-pharmacological intervention description and reported that only 39% were adequately described.

Rates of replicability of interventions vary considerably in the published literature depending on the complexity of the treatment and the assessment criteria. For example, three studies

assessed compliance with item four of CONSORT in published research in the areas of weight loss[5], brain tumours[6] and Hodgkin's lymphoma[7]. Item four of CONSORT specifically asks for precise details concerning treatments intended for all groups and how and when they were administered. These studies reported that over 90% of study findings were replicable. In contrast, however, one study assessed whether there was sufficient information on what happens before, during and after treatment for back pain, and revealed that only 13% of trials were replicable[8].

The NIHR Health Technology Assessment (HTA) Programme commissions and funds primary research and evidence synthesis on the effectiveness, costs and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the NHS. It aspires to enable all funded projects to complete and publish in the programme's own Journal '*Health Technology Assessment*', freely available on the programme's website (www.hta.ac.uk). Reports published in the Journal series are peer reviewed, are in the public domain and contain a full record of the study. Unlike typical peer reviewed journals, there are no word or size limitations for the full report and unlimited appendices, thus enabling more detail to be included in the publication; an average report is approximately 50,000 in length. Given the importance of complete and replicable reporting of findings and the opportunities the NIHR HTA Journal presents, this study aimed to assess whether randomised controlled trials (RCTs) with single trials published in *Health Technology Assessment* were described in sufficient detail.

Methods:

Data source

All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion in the study. Of the 109 reports publishing in this time-period, 11 were excluded as they reported more than one RCT within a single HTA Journal. Ninety-eight single trial RCTs were therefore included in the study.

Piloting the checklist

Five NIHR HTA funded RCTs were selected to represent a range of interventions (surgery, psychology, devices and pharmaceutical) to pilot the checklist initially developed by *Schroter et al*[4]. The checklist was applied independently by three assessors to assess the level of agreement between assessments. A kappa score was produced for each individual trial (range 0.15 – 0.7) with an average of 0.225 across all 5 trials. Disagreement was due to differing interpretations of the checklist questions. The initial checklist was modified to separate out two of the questions into their individual components. In the published checklist the recipient question stated 'Is it clear who is receiving the intervention?' and 'Do you know all that you need to about the patients? (e.g. which drugs they are taking, what they were told, etc)?'. We felt that this required several pieces of information for a single question and therefore we separated question 2 into the three components, as shown in table 1. Similarly the materials question, 'Are the physical or informational materials used adequately described?' was separated into the two components shown in question 7 of table 1. Additionally, the assessors discussed the type and level of information expected to be present in order to answer a question as complete. The modified checklist was applied by the three assessors to a further five NIHR HTA funded RCTs, resulting in higher levels of agreement (kappa scores for each report ranged from 0.3 – 0.7, with an average of 0.6 for all trials).

The main study

The final modified checklist was applied to a wider sample of NIHR HTA funded RCTs. All RCTs published in the NIHR HTA Journal (from January 1999 until March 2011) were selected for inclusion with in the study. One checklist was completed for the intervention group of each trial published. Each item in the checklist was answered by either a yes, no or not applicable response. We did not apply the full checklist to the Control Group, but, unlike the published checklist (Schroter et al[4]), we did make a general assessment as the to completeness of control group information within question nine of the checklist. However, responses to this question were not on a detailed assessment of all components of the

control group, unlike the intervention group itself. Question eight summarises whether there are any aspects of the intervention missing based on the responses to the previous seven questions.

Data quality

Each trial was assessed independently by two assessors. 15% of published reports (15/98) were discussed due to disagreements of the scoring mainly around checklist item seven. All disagreements were discussed by the team and were resolved by consensus.

Three assessors carried out the assessments. Each trial was allocated to two assessors who independently applied the criteria. None of the assessors had medical or clinical experience; however have higher health degrees and work full time in health research and in evaluating clinical research. All NIHR HTA reports were examined by using a stabilised process; initially scanning the executive summary, followed by the methods, using key word search terms to scan the whole document and appendices, and finally undertaking a detailed reading of the entire report if relevant information could not be found.

Data Analysis

The checklist for each trial was completed using an electronic, stand-alone Access database. All three assessors completed checklists were then merged and exported into Excel and IBM SPSS version 19 for data analysis. IBM SPSS software was used to conduct all descriptive and inferential analyses. The chi-square test was used for all comparisons (statistically significant at P <0.05). If any cell had an expected count less than 5, the Fisher's exact test was used.

Results:

The modified checklist was applied to 98 RCTs published the HTA Journal series from January 1999 until March 2011. The interventions within each published trial were classified by the following intervention types; pharmaceutical, radiotherapy, surgery, diagnostic, education and training, service delivery, psychological, vaccines and biological, devices, physical therapy, exercise, complementary therapy, mixed or complex and other. The intervention classification was provided by *Schroter et al*[4] as part of the original checklist. Table 2 shows the number of trials within the journal series for each intervention by type.

Applying the modified checklist to NIHR HTA funded RCTs revealed that components of the intervention description were missing in 68 of the 98 reports (missing 69.4%). Table 3 contains selected examples against six of the items in the checklist to illustrate both complete and poorly described interventions.

Intervention descriptions were therefore complete in 30.6% of reports. Certain criteria had high levels of completeness, such as baseline characteristics (94.9%) and descriptions of settings (91.8%), which were complete for over 90% of reports. However, other criteria were notably less complete, particularly patient information with only 58.2% having an adequate description (Table 4).

Differences in completion rates were noted between the fourteen types of interventions. For example, descriptions of interventions were more complete for drug interventions than non-drug interventions with 33.3% and 30.6% levels of completeness respectively. The chi-square test showed that this difference was not statistically significant (p=0.77). Furthermore, this was not the case with certain criteria, such as baseline characteristics (drugs 93.3%, non-drugs 94.4%) and provider information (drugs 73.3, non-drugs 77.8%) where levels of completeness were higher in non-drug trials than drug interventions.

Descriptions of interventions were found to be least complete for psychological interventions with only 27.3% of RCTs in this area being complete. The chi-squared test revealed that this difference was not statistically significant when compared with drug interventions (p=1.00). Again, there were a few occasions where certain criteria had the highest levels of completeness of all intervention types, in particular with baseline characteristics and provider information with 100% and 90.9% of completeness respectively (Table 4).

The modified checklist included a question around the completeness of the control group. This was not a detailed evaluation of all the components of the control group but a broad assessment of whether the description appeared to be complete or not. Given the

interpretative nature of this question, control group information were not included with the full data. The data revealed that 51% of RCTs had complete descriptions of control groups.

Discussion:

Statement of principle findings

This study has revealed that 30.6% (30/98) of studies with a single trial published in the HTA Journal have a full description of the intervention. The interventions described in published RCTs performed well against certain criteria, such as baseline characteristics (with 95% having an adequate description), but less well on other criteria, such as patient information (with 58% having an adequate description). Drug trials were slightly more complete than non-drug trials and psychological interventions with 33.3% of journals having a complete intervention description, although these differences were not statistically significant.

Strengths and weaknesses

The strengths of this study are that externally generated and tested criteria were applied to evaluate the effectiveness of intervention descriptions in NIHR HTA Programme funded RCTs. However, there were limitations. First, none of the assessors applying the criteria were medically trained, however, assessors were not commenting on the suitability of an intervention for use in practice but whether aspects of the description that would be required for use in practice were present. There is a possibility that someone with medical training would score the projects differently. In previous work the authors have been medically trained. Second, authors of the reports were not contacted to provide additional information beyond that provided in the publication. Previous studies have demonstrated that contacting the research teams or additional searches for intervention details does increase the completeness of intervention descriptions[3,9]. However it is questionable whether having to undertake additional searches outside the publication effectively enhances the ease of replicating study findings.

A limitation of the checklist used is the type of data being collected. ~~third limitation concerned the type of data being collected.~~ Whilst all the criteria are dichotomous (in that they are all yes / no answers), the justification behind this categorisation has different degrees of interpretation. This could have resulted in overly-harsh assessments of completeness for certain criteria. For example, the recipient criterion is clear (are inclusion/exclusion criteria present) whilst greater interpretation is required for the materials criterion which requires the assessor to determine if the description of the physical materials is adequate and therefore open to interpretation. Certainly the completion rate for materials was among the lowest across all studies with 58% and 69% completion rates for informational and physical materials respectively. By using this checklist we were able to suggest further refinements to the criteria used within it, such as separating out the recipient criteria and the material criteria.

A further limitation of the study was that the checklist was not fully applied to the control group of the published trials. This would have provided a more complete picture of how well controls are described within a study. Another limitation was that the number of journals assessed for completeness was very small for certain assessments (for example only 11 / 98 journals reported psychological interventions). It is possible therefore, that certain findings of completeness rate occurred by chance.

Meanings of the study

It is tempting to make comparisons with others studies assessing the usability of intervention descriptions. In particular, *Glasziou et al*[3] reported that 41/80 (51%) of published reports of single randomised trials and systematic reviews in popular journals were complete compared with 30/98 (30.6%) completeness of NIHR HTA funded RCT trials. Similarly, interventions in NIHR HTA reports appeared to be described less well than the 51 trials published in the BMJ assessed by *Schroter et al.* where 43% (22/51) of papers were considered to be of sufficient description to allow replication[4]. Whilst these comparisons are interesting, it is important to note that it is not possible to make any meaningful comparison on the relative performance of each output, as the *Glasziou et al*[3] study looked at journal articles and we looked at the HTA journal series which are aimed at different audiences and the questionnaire used was different between the studies. This is because the nature of outputs varies considerably between studies as does the assessment criteria. It is notable, for example, that *Schroter et al*[4] used eight indicators (seven main checklist items and a global completeness eighth item) in their checklist, compared with the twelve criteria used in this study.

However, this study does reflect findings from similar studies conducted elsewhere. For example, the criteria highlighted as being particularly poorly described in Schroter's study were physical / informational materials, which reflected findings in this study where patient information and physical materials were also lacking in completeness. Similarly the fact that NIHR HTA Programme funded drug interventions were typically better described than non-drug interventions reflected findings in *Glasziou et al*[3]. where over 60% of reports on drug treatments were initially deemed to be complete compared with just under 30% of non- drug treatments.

In addition to the more detailed guidance provided to authors, the HTA Journal requests that authors of RCTs include the headings set out in the revised CONSORT checklist and flowchart and provide details of CONSORT in its guidance for authors. Item five of the

CONSORT statement says 'The interventions for each group with sufficient details to allow replication, including how and when they were actually administered' and there are extensions of the CONSORT statement to address the additional complexity around the reporting of non-pharmacological interventions. The CONSORT extensions are not currently a requirement for non-pharmacological studies but as these extensions are more widely requested, it is hopeful that the reporting of interventions will improve and be fully described.

A number of studies have investigated the completeness of intervention descriptions in a single disease area by assessing compliance of RCTs with the intervention item (item 4) of the CONSORT statement[5,6,7]. Whilst these studies reported that over 90% of study findings were replicable, it is likely that this is an over estimation as they do not assess the question of whether there was enough information to allow replication. In contrast, however, one study assessed whether there was sufficient information on what happens before, during and after treatment for back pain, and revealed that only 13% of trials were replicable[8].

Understanding the extent to which interventions in published studies are described sufficiently to inform clinical decision making is a key concern in the adding value in research agenda. As Chalmers and Glasziou have suggested, poorly described interventions form one of the four main pillars of research waste[1]. The criteria identified by *Schroter et al*[4] and developed in this study are helpful in highlighting specific areas of where intervention descriptions can be improved.

Future research

Several areas for further research are indicated by this study. Further testing on the criteria can be undertaken to assess the repeatability of the criteria. For example, the reports sampled in this study could be reassessed by someone with clinical experience to assess the level of agreement. Alternatively, Glasziou's selected papers in his original study could be assessed by non-clinical teams to examine the level of agreement. The checklist has only been applied to single trial studies; future research into the applicability of it for multi-trial studies should be investigated.

The characterisation of the control group is a key area for future research, as research involving trials to date has focused on the description of interventions with a treatment group, however the detail of the control arm is equally important as in many cases the control arm is often described as 'usual care' but this does not take into account variations by centre[10]. A recent paper reported on the development of a tool for extraction of data in systematic reviews and includes an element on intervention design[11]. The tool has been applied to both

the intervention and control groups of systematic reviews. The applicability of the tool across primary research could be investigated and used to further strengthen the checklist that we have used.

Ensuring the replicability of study findings is an essential part of adding value in research. It is important for health research publishers to be transparent in the usability of study reports and areas of improvement. This study applied a checklist that can be used to indicate where the descriptions of interventions can be improved to enhance replication in clinical practice. Serious consideration should be given on how this might be used to improve intervention reporting in the future. The results of this study have been shared with the editorial Board of the HTA Journal to investigate how interventions can be better reported within the journal series.

References

1. Chalmers I,Glasziou P. Avoidable waste in the production and reporting of research evidence. *The Lancet* 2009;**374**:86-9
2. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ* 2010; 341.
3. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008; 336(7659):1472-1474.
4. Schroter S, Glasziou P, Heneghan C. Quality of description of treatments: a review of published randomised controlled trials. BMJ Open 2012; 2.
5. Thabane L, Chu R, Cuddy K, Douketis J. What is the quality of reporting in weight loss intervention studies? A systematic review of randomized controlled trials. [Review] [17 refs]. *International Journal of Obesity* 2007; 31(10):1554-1559
6. Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in Hodgkin's lymphoma in biomedial journals. *J Natl Cancer Inst* 2006; 98:620-625.
7. Lai R, Chu R, Fraumeni M, et al. Quality of ransomized controlled trials reporting in the primary treamt of brain tumors. *J Clin Oncol* 2006; 24:1136-1144.
8. Glenton C, Underland V, Kho M, Pennick V, Oxman AD. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. *Journal of Clinical Epidemiology* 2006; 59(8):770-778
9. Hoffmann T, Erueti C, Glasziou P. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ* 2013; 347:f3755
10. Cook A, Douet L and Boutron I. Descriptions of non-pharmacological interventions in clinical trials BMJ 2013;347:f5212
11. Montgomerya P, Underhill K, Gardnera F, Operarioc D, Mayo-Wilson E. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology* 2013 66(8): 874-882

Table 1: Replicability Criteria for Interventions, developed from an initial design by Schroter et al[3]

Checklist criteria		Descriptor of criteria where appropriate
1. Setting	Is it clear where the intervention was delivered?	
2a. Recipient – inclusion	Is it clear who is receiving the intervention? – Inclusion criteria	Clear inclusion criteria in the journal
2b. Recipient – exclusion	Is it clear who is receiving the intervention? – Exclusion criteria	Clear exclusion criteria in journal
2c. Recipient – baseline characteristics	Do you know all that you need to about the patients? (e.g. which drugs they are taking, what they were told, etc)? If No, what further information do you require?	Baseline characteristics of participants provided in journal
3. Provider	Is it clear who delivered the intervention?	
4. Procedure	Is the procedure (including the sequencing of the technique) of the intervention sufficiently clear to allow replication?	Top level overview of the intervention. Eg Drug X for X days at X dose. Or X sessions lasting X minutes, for X weeks/ months
5. Intensity	Is the dose/duration of individual sessions of the intervention clear?	Dose, length of session
6. Schedule	Is the schedule (interval, frequency, duration, or timing) of the intervention clear?	Frequency of intervention, length of session
7a. Materials – physical	Are the physical materials used adequately described?	Physical materials eg Description of splint used. If either are no, it is a no overall
7b. Materials - informational	Are the informational materials used adequately described?	Information provided to the patients eg consent forms etc
8. Missing	Is the description of the intervention complete? If No, what is missing?	
9. Control	Is it clear what the control group received during the study?	

Table 2: Intervention type of NIHR HTA funded trial RCTs included in the study

Type of intervention	№ (%)
Drug	15 (15.3)
Radiotherapy	1 (1.0)
Surgery	9 (9.2)
Diagnostic	8 (8.2)
Education and training	3 (3.1)
Service delivery	19 (19.4)
Psychological therapies	11 (11.2)
Vaccines and biologicals	3 (3.1)
Devices	12 (12.2)
Physical therapies	7 (7.1)
Exercise	1 (1.0)
Complementary therapies	2 (2.0)
Mixed or complex	6 (6.1)
Other*	1 (1.0)
Total	**98**

*Other refers to an intervention using larval therapy

Table 3: Examples of poor reporting of intervention elements within the HTA journal series, taken verbatim form the journal.

Checklist item	Examples of poor reporting	Reason why rated as incomplete	Examples of good reporting	Reason why rated complete
Inclusion criteria	'….patient identification was retrospective. Searches were conducted on practice databases using either repeat prescriptions alone or repeat prescriptions plus diagnostic terms… GPs then sent letters to suitable patients, providing information about the trial'	No details given about the searches and the criteria patients were screened with.	'Inclusion criteria for trial patients were: -Diagnosed with idiopathic arthritides of childhood with onset before their 16th birthday for more than 3 months. - Aged 4–19 years inclusive. - Stable on medication. - At least one active joint, core set criteria 1.56 -At least two out of any five of the remaining core set criteria below. -The physician global assessment of disease activity >10 mm on a 100-mm visual analogue scale (VAS). - The parent global assessment of well-being >10 mm on a 100-mm VAS. -Childhood Health Assessment Questionnaire scores >0. - More than one joint with limited range of motion (joint motion reduced by at least 5° from normative range for age58). -An elevated erythrocyte sedimentation rate (ESR) (>5 mmHg in children and >10 mmHg in adolescents)'	Very detailed patient criteria listed
Exclusion criteria	'GPs were given a ringbinder file with information and instructions about the trial and, within each, a number of recruitment packs. The packs contained the paperwork required to complete the recruitment of each patient, this was: a reminder of the inclusion/exclusion criteria for the study….'	No details given about the exclusion criteria	'Reasons for Exclusion (Yes/No) - BMI > 40 kg/m2 - Barrett's oesophagus (≥3cm) - Paraoesophageal hernia - Oesophageal strictures - One type of management is clinically indicated for another reason'	Detailed patient exclusion criteria listed
Provider	'All services had staff who were trained and experienced in family therapy, but not necessarily family interventions specifically for eating disorders'	No details about the staff providing the interventions or the training they received.	'Eight counsellors (six females and two males) took part in the trial (one worked at two practices) and all were BAC accredited or eligible for BAC accreditation; they were highly trained and had considerable experience of counselling in a general practice setting' (There are details about each counsellors age, qualifications and experience are provided)	States who delivered the intervention and their training
Procedure	'Generally home-based rehabilitation services provide, as a minimum, physiotherapy and occupational therapy in the patient's own home. Services can be specialised (e.g. in stroke rehabilitation) or be provided for patients with a range of disabilities'	No details about the services provided to patients and variation between centres	'The content of the CBT programme included (complete course description contained within an appendix): - Elucidation of core beliefs regarding their illness and its management. - Monitoring of activity levels and introduction of appropriate timetable. - Introduction to exercises designed to increase general level of fitness, balance and confidence in exercise. A range of aerobic, strength, balance and stretching exercises were taught. - Behavioural modification of sleep patterns. - Mood management advice. - Goal setting'	Key aspects of the intervention summarised in the text and a full description of the intervention is detailed in the appendices.
Intensity and Schedule	'Patients come to the day hospital where the rehabilitation service is provided for a full or half day. Usually ambulance transport is provided to bring patients into the service and return them home after a session'	No details of the length or number of sessions	'Psychological treatment was based on existing protocols (references included) and distributed over six 50-minute sessions, with printed information sheets provided after each session'	The length and number of sessions is included as well as the details of each session.
Materials – physical	'The acupuncture point prescriptions used were individualised to each patient and were at the discretion of the acupuncturist'	The prescriptions used are not detailed.	'- 500 mg oral oxytetracycline (non-proprietary) b.d. + topical vehicle control b.d. -100 mg oral Minocin MR minocycline) o.d. + topical vehicle control b.d. -topical Panoxyl Aquagel (5% benzoyl peroxide) b.d. + oral placebo o.d. This was designated as the active comparator group, as benzoyl peroxide was the leading and most established topical treatment for acne when the protocol was written. -topical Benzamycin (3% erythromycin + 5% benzoyl peroxide) b.d. + oral placebo o.d. (referred to as ery. + BP bd) -topical Stiemycin (2% erythromycin) o.d. +topical Panoxyl Aquagel (5% benzoyl peroxide)o.d. + oral placebo o.d. (referred to as ery. od+ BP od)'	Each of the treatments prescribed is clearly defined

Table 4: Completion rates of NIHR HTA reports by intervention type

Description criteria	Drugs No. Trials with complete intervention description (%)	CI*	Non-Drugs No. Trials with complete intervention description (%)	CI*	Psychological No. Trials with complete intervention description (%)	CI*	All No. Trials with complete intervention description (%)
Number journals	15 (15.3)		72 (73.5)		11 (11.2)		98 (100)
Setting	15 (100)	0.80-1	66 (91.7)	0.83-0.96	9 (81.8)	0.52-0.95	90 (91.8)
Inclusion criteria	15 (100)	0.80-1	63 (87.5)	0.78-0.93	10 (90.9)	0.62-0.98	88 (89.8)
Exclusion criteria	14 (93.3)	0.70-0.99	55 (76.4)	0.65-0.85	10 (90.9)	0.62-0.98	79 (80.6)
Baseline characteristics	14 (93.3)	0.70-0.99	68 (94.4)	0.87-0.98	11 (100)	0.74-1	93 (94.9)
Provider	11 (73.3)	0.48-0.89	56 (77.8)	0.67-0.86	10 (90.9)	0.62-0.98	77 (78.6)
Procedure	14 (93.3)	0.70-0.99	57 (79.2)	0.68-0.87	9 (81.8)	0.52-0.95	80 (81.6)
Intensity	13 (86.7)	0.62-0.96	63 (87.5)	0.78-0.93	9 (81.8)	0.52-0.95	85 (86.8)
Schedule	13 (86.7)	0.62-0.96	59 (81.9)	0.71-0.89	9 (81.8)	0.52-0.95	81 (82.7)
Patient information	9 (60.0)	0.36-0.80	42 (58.3)	0.47-0.69	6 (54.5)	0.28-0.79	57 (58.2)
Physical materials	10 (66.7)	0.42-0.85	52 (72.3)	0.61-0.81	6 (54.5)	0.28-0.79	68 (69.4)
Intervention description complete overall	5 (33.3)	0.15-0.58	22 (30.6)	0.21-0.42	3 (27.3)	0.10-0.57	30 (30.6)

**Please note: some criteria are not applicable therefore denominator less than total number of journals.*95% confidence interval: no continuity correction*