

BMJ Open The completeness of intervention descriptions in published National Institute of Health Research HTA-funded trials: a cross-sectional study

Lisa Douet,¹ Ruairidh Milne,² Sydney Anstee,¹ Fay Habens,¹ Amanda Young,¹ David Wright¹

To cite: Douet L, Milne R, Anstee S, *et al.* The completeness of intervention descriptions in published National Institute of Health Research HTA-funded trials: a cross-sectional study. *BMJ Open* 2014;**4**:e003713. doi:10.1136/bmjopen-2013-003713

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2013-003713>).

Received 2 August 2013
Revised 28 November 2013
Accepted 29 November 2013



CrossMark

¹National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre (NETSCC), University of Southampton, Southampton, UK

²Wessex Institute, University of Southampton, Southampton, UK

Correspondence to

Dr Lisa Douet;
l.douet@southampton.ac.uk

ABSTRACT

Objectives: The objective of this study was to assess whether National Institute of Health Research (NIHR) Health Technology Assessment (HTA)-funded randomised controlled trials (RCTs) published in the *HTA* journal were described in sufficient detail to replicate in practice.

Setting: RCTs published in the *HTA* journal.

Participants: 98 RCTs published in the *HTA* journal up to March 2011. Completeness of the intervention description was assessed independently by two researchers using a checklist, which included assessments of participants, intensity, schedule, materials and settings. Disagreements in scoring were discussed in the team; differences were then explored and resolved.

Primary and secondary outcome measures:

Proportion of trials rated as having a complete description of the intervention (primary outcome measure). The proportion of drug trials versus psychological and non-drug trials rated as having a complete description of the intervention (secondary outcome measures).

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.

Strengths and limitations of this study

- An externally produced checklist was applied to all randomised controlled trials published in the National Institute of Health Research (NIHR) *Health Technology Assessment* journal series.
- The sample size for a number of assessments is very small.
- The checklist was only applied to the intervention arm of the trial, in the future it would be important to apply the checklist fully to both the control and intervention arm of trials.

INTRODUCTION

A recent publication by Chalmers and Glasziou¹ 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.

Several studies have specifically assessed the waste area for ensuring that the 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. While 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 the 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 the next day.³ Of these 80 published reports, 41 (51%) had elements of the intervention missing, particularly descriptions of process and information on handouts 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*⁴ developed, piloted and applied a checklist designed to assess the replicability of published treatment decisions to 51 trials published in the *BMJ*. 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 the 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 study⁵ 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 4 of CONSORT in published research in the areas of weight loss,⁶ brain tumours⁷ and Hodgkin's lymphoma.⁸ Item 4 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 the trials were replicable.⁹

The National Institute of Health Research (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 *HTA*, freely available on the programme's website (<http://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 present, this study aimed to

assess whether randomised controlled trials (RCTs) with single trials published in *HTA* 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 published in this time period, 11 were excluded as they had 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 κ score was produced for each individual trial (range 0.15–0.7) with an average of 0.225 across all five trials. Disagreement was due to differing interpretations of the checklist questions. The initial checklist was modified to separate 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? (eg, 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 material question, 'Are the physical or informational materials used adequately described?' was separated into two components, which is shown in question 7 of table 1. In addition, 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 (κ scores for each report ranged from 0.3 to 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 to 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,⁴ we did make a general assessment as to the completeness of control group information within question 9 of the checklist. However, responses to this question were not on a detailed assessment of all components of the

Table 1 Replicability criteria for interventions, developed from an initial design by Schroter *et al*^β

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? (eg, 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?	
5. Intensity Is the dose/duration of individual sessions of the intervention clear?	
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?	
7b. Materials—informational Are the informational materials used adequately described?	
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?	

control group, unlike the intervention group itself. Question 8 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. Fifteen per cent of the published reports (15/98) were discussed due to disagreements of the scoring mainly around checklist item 7. 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, they 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. The checklists

completed by all three assessors were then merged and exported into Excel and IBM SPSS V.19 for data analysis. IBM SPSS software was used to conduct all descriptive and inferential analyses. The χ^2 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 by 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*[†] 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 complete and poorly described interventions.

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

Type of intervention	N (%)
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.
HTA, Health Technology Assessment; RCT, randomised controlled trial.

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 14 types of interventions. For example, descriptions of interventions were more complete for drug interventions than for non-drug interventions with 33.3% and 30.6% levels of completeness, respectively. The χ^2 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 in 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 χ^2 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 about 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 principal 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 the 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 were discussing on whether the aspects of the description would be required for use in practice. There is a possibility that someone with medical training would score the projects differently. In a previous work, the authors have been medically trained. Second, the 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 5} 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. While 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) while greater interpretation is required for the materials criterion which requires the assessor to determine whether 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.

Table 3 Examples of poor reporting of intervention elements within the HTA journal series, taken verbatim from 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: <ul style="list-style-type: none"> ▶ 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 of any five of the remaining core set criteria below ▶ The physician global assessment of disease activity >10 mm on a 100-mm 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 ESR (>5 mm Hg in children and >10 mm Hg 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) <ul style="list-style-type: none"> ▶ BMI >40 kg/m² ▶ Barrett's oesophagus (≥3 cm) ▶ 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 BACP accredited or eligible for BACP 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	No details about the services provided to	The content of the CBT programme included (complete course description contained within an appendix):	Key aspects of the intervention summarised in the text and a full description of the

Continued

Table 3 Continued

Checklist item	Examples of poor reporting	Reason why rated as incomplete	Examples of good reporting	Reason why rated complete
	the patient's own home. Services can be specialised (eg, in stroke rehabilitation) or be provided for patients with a range of disabilities'	patients and variation between centres	<ul style="list-style-type: none"> ▶ 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 	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	<ul style="list-style-type: none"> ▶ 500 mg oral oxytetracycline (non-proprietary) twice daily+topical vehicle control twice daily ▶ 100 mg oral Minocin MR (minocycline) once daily+topical vehicle control twice daily ▶ Topical Panoxyl Aquagel (5% benzoyl peroxide) twice daily+oral placebo once daily. 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) twice daily+oral placebo once daily (referred to as ery.+BP twice daily) ▶ Topical Stiemycin (2% erythromycin) once daily+topical Panoxyl Aquagel (5% benzoyl peroxide) once daily+oral placebo once daily (referred to as ery. once daily+ BP once daily) 	Each of the treatments prescribed is clearly defined

BACP, British Association for Counselling & Psychotherapy; BP, blood pressure; CBT, cognitive behavioural therapy; ESR, erythrocyte sedimentation rate; GP, general practitioner; HTA, Health Technology Assessment; VAS, visual analogue scale.

Table 4 Completion rates of NIHR HTA reports by intervention type

Description criteria	Drugs Number of trials with complete intervention		Non-drugs Number of trials with complete intervention		Psychological Number of trials with complete intervention		All Number of trials with complete intervention	
	description (%)	95% CI	description (%)	95% CI	description (%)	95% CI	description (%)	95% CI
Number journals	15 (15.3)		72 (73.5)		11 (11.2)		98 (100)	
Setting	15 (100)	0.80 to 1	66 (91.7)	0.83 to 0.96	9 (81.8)	0.83 to 0.96	90 (91.8)	0.52 to 0.95
Inclusion criteria	15 (100)	0.80 to 1	63 (87.5)	0.78 to 0.93	10 (90.9)	0.78 to 0.93	88 (89.8)	0.62 to 0.98
Exclusion criteria	14 (93.3)	0.70 to 0.99	55 (76.4)	0.65 to 0.85	10 (90.9)	0.65 to 0.85	79 (80.6)	0.62 to 0.98
Baseline characteristics	14 (93.3)	0.70 to 0.99	68 (94.4)	0.87 to 0.98	11 (100)	0.87 to 0.98	93 (94.9)	0.74 to 1
Provider	11 (73.3)	0.48 to 0.89	56 (77.8)	0.67 to 0.86	10 (90.9)	0.67 to 0.86	77 (78.6)	0.62 to 0.98
Procedure	14 (93.3)	0.70 to 0.99	57 (79.2)	0.68 to 0.87	9 (81.8)	0.68 to 0.87	80 (81.6)	0.52 to 0.95
Intensity	13 (86.7)	0.62 to 0.96	63 (87.5)	0.78 to 0.93	9 (81.8)	0.78 to 0.93	85 (86.8)	0.52 to 0.95
Schedule	13 (86.7)	0.62 to 0.96	59 (81.9)	0.71 to 0.89	9 (81.8)	0.71 to 0.89	81 (82.7)	0.52 to 0.95
Patient information	9 (60.0)	0.36 to 0.80	42 (58.3)	0.47 to 0.69	6 (54.5)	0.47 to 0.69	57 (58.2)	0.28 to 0.79
Physical materials	10 (66.7)	0.42 to 0.85	52 (72.3)	0.61 to 0.81	6 (54.5)	0.61 to 0.81	68 (69.4)	0.28 to 0.79
Intervention description	5 (33.3)	0.15 to 0.58	22 (30.6)	0.21 to 0.42	3 (27.3)	0.21 to 0.42	30 (30.6)	0.10 to 0.57
complete overall								

Some criteria are not applicable, therefore denominator less than total number of journals. 95% CI: no continuity correction. HTA, Health Technology Assessment.

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 (eg, only 11/98 journals reported psychological interventions). Therefore, it is possible 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*⁸ reported that 41/80 (51%) of the 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*⁴ where 43% (22/51) of the articles were considered to be of with sufficient description to allow replication. While 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*⁸ 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 (7 main checklist items and a global completeness eighth item) in their checklist, compared with the 12 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 the findings in *Glasziou et al*⁸ 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 5 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 the compliance of RCTs with the intervention item (item 4) of the CONSORT statement.^{6–8} While 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 the trials were replicable.⁸

Understanding the extent to which interventions in the published studies are described sufficiently to inform clinical decision-making is a key concern in adding value in research agenda. As Chalmers and Glasziou¹ have suggested, poorly described interventions form one of the four main pillars of research waste. The criteria identified by Schroter *et al*⁴ and developed in this study are helpful in highlighting the 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 multitrail 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 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 a journal series.

Acknowledgements The authors 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.

Contributors 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 the guarantor.

Funding 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.

Competing interests 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* (2011–2012).

Ethics approval None.

Provenance and peer review Not commissioned; externally peer reviewed.

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

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;374:86–9.
- Glasziou P, Chalmers I, Altman DG, *et al*. Taking healthcare interventions from trial to practice. *BMJ* 2010;341:c3852.
- Glasziou P, Meats E, Heneghan C, *et al*. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008;336:1472–4.
- Schroter S, Glasziou P, Heneghan C. Quality of description of treatments: a review of published randomised controlled trials. *BMJ Open* 2012;2:e001978.
- Hoffmann TC, Eructi C, Glasziou PP. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ* 2013;347:f3755.
- Thabane L, Chu R, Cuddy K, *et al*. What is the quality of reporting in weight loss intervention studies? A systematic review of randomized controlled trials. [Review] [17 refs]. *Int J Obes* 2007;31:1554–9.
- Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in Hodgkin's lymphoma in biomedical journals. *J Natl Cancer Inst* 2006;98:620–5.
- Lai R, Chu R, Fraumeni M, *et al*. Quality of randomized controlled trials reporting in the primary treatment of brain tumors. *J Clin Oncol* 2006;24:1136–44.
- Glenton C, Underland V, Kho M, *et al*. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. *J Clin Epidemiol* 2006;59:770–8.
- Cook A, Douet L, Boutron I. Descriptions of non-pharmacological interventions in clinical trials. *BMJ* 2013;347:f5212.
- Montgomery P, Underhill K, Gardner F, *et al*. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *J Clin Epidemiol* 2013;66:874–82.