Physiotherapy 102 (2016) 121–126

Systematic review

How completely are physiotherapy interventions described in reports of randomised trials?☆

Tiê P. Yamato [a,*], Chris G. Maher [a], Bruno T. Saragiotto [a], Tammy C. Hoffmann [b], Anne M. Moseley [a]

[a] *Musculoskeletal Division, The George Institute for Global Health, Sydney Medical School, The University of Sydney, Australia*
[b] *Centre for Research in Evidence Based Practice, Faculty of Health Sciences and Medicine, Bond University, Queensland, Australia*

Abstract

Background Incomplete descriptions of interventions are a common problem in reports of randomised controlled trials. To date no study has evaluated the completeness of the descriptions of physiotherapy interventions.
Objectives To evaluate the completeness of the descriptions of physiotherapy interventions in a random sample of reports of randomised controlled trials (RCTs).
Data sources A random sample of 200 reports of RCTs from the PEDro database.
Study selection or eligibility criteria We included full text papers, written in English, and reporting trials with two arms. We included trials evaluating any type of physiotherapy interventions and subdisciplines.
Data extraction and data synthesis The methodological quality was evaluated using the PEDro scale and completeness of intervention description using the Template for Intervention Description and Replication (TIDieR) checklist. The proportion and 95% confidence interval were calculated for intervention and control groups, and used to present the relationship between completeness and methodological quality, and subdisciplines.
Results Completeness of intervention reporting in physiotherapy RCTs was poor. For intervention groups, 46 (23%) trials did not describe at least half of the items. Reporting was worse for control groups, 149 (75%) trials described less than half of the items. There was no clear difference in the completeness across subdisciplines or methodological quality.
Limitations Our sample were restricted to trials published in English in 2013.
Conclusion and implications of key findings Descriptions of interventions in physiotherapy RCTs are typically incomplete. Authors and journals should aim for more complete descriptions of interventions in physiotherapy trials.
© 2016 Chartered Society of Physiotherapy. Published by Elsevier Ltd. All rights reserved.

Keywords: Clinical trial; Research design; Physical therapy specialty

Introduction

For clinicians managing patients within an evidence-based practice framework, two key elements are required before management can commence [1]. The first one is to locate high-quality clinical research demonstrating the efficacy of the intervention, and the second is to have an adequate description of the intervention procedure(s) so that it can be replicated with the patient in question [1].

While there is guidance on how to develop and evaluate complex interventions [2] and an extensive body of literature on the quality of research about intervention effectiveness, research into the adequacy of the description of the interventions in trials has had little attention [1]. Previous studies have revealed that incomplete description of interventions is a common problem in reports of randomised controlled trials

☆ A full reference list of the papers analysed in this trial is available from the authors.

* Correspondence: Level 13/321 Kent Street, 2000 Sydney, NSW, Australia. Tel.: +61 0404 054 818.
E-mail address: tiparma@gmail.com (T.P. Yamato).

http://dx.doi.org/10.1016/j.physio.2016.03.001
0031-9406/© 2016 Chartered Society of Physiotherapy. Published by Elsevier Ltd. All rights reserved.

of medical interventions [3,4]. This may affect the replication of an intervention as for example providers may not understand the content of the intervention or the equipment and skills required for effective delivery [5–7]. Clear description of an intervention facilitates understanding of its components and the relationship between them, and aids outcome interpretation and process evaluation, particularly in complex interventions [8].

Given the importance of adequate reporting of interventions in publications and the lack of existing guidance, the Template for Intervention Description and Replication (TIDieR) was developed [9]. TIDieR has the objective of providing further guidance on complete intervention reporting and is an extension to the CONSORT 2010 [10] and SPIRIT 2013 [11] statements. It is a 12-item checklist that includes the minimum recommended items for describing interventions and was developed following the recommended procedures for developing a reporting guideline [12].

The completeness of reporting of interventions has not been specifically investigated in published trials of physiotherapy interventions. Therefore, this study aimed to evaluate the completeness of the descriptions of physiotherapy interventions in a random sample of reports of randomised controlled trials.

Methods

Type of studies

From the Physiotherapy Evidence Database (PEDro; www.pedro.org.au), we extracted a random sample of 200 reports of randomised controlled trials published in 2013 by using the random number function in Microsoft Excel software (Microsoft Office 2007, Microsoft Corporation, Redmond, Washington). We included full text papers, written in English, and reporting trials with two arms (where an active intervention was compared with a less active intervention, no treatment or a placebo). Types of physiotherapy treatment or physiotherapy subdiscipline were not inclusion criteria.

The bibliometric data and the subdiscipline of physiotherapy (sports, paediatrics, orthopaedics, oncology, neurology, musculoskeletal, gerontology, ergonomic and occupational health, continence and women's health, cardiothoracic and other) were downloaded from PEDro. Full-text copies of the included studies were acquired and, if mentioned in the trial report, we also obtained any online supplementary material describing the interventions.

Assessment of methodological quality

The methodological quality of included studies was evaluated using the 10-point PEDro scale, with the total PEDro score for each trial report downloaded from the PEDro database. The PEDro scale has been found to have acceptable reliability [13] and validity [14,15].

Assessment of intervention descriptions

We assessed the completeness of the description of the interventions using the TIDieR checklist [9]. We used the information contained in the published trial report plus any online supplementary materials that were referred to in the trial report. Each item was scored as 'yes', 'not applicable' or 'no' for both the intervention and control groups in each included study. The 'not applicable' code was only used for control groups where no treatment was administered (e.g. wait list or no treatment). A description of the 12 items of TIDieR is presented in Appendix 1.

A customised Excel spreadsheet was used for the TIDieR [9] checklist for each included study. Two independent reviewers rated the reporting of both the intervention and control conditions of the included studies using the TIDieR checklist for the first 100 included studies. For the second 100 studies, TIDieR items with sufficient inter-rater reliability (Prevalence and Bias Adjusted Kappa ≥ 0.60) were scored by a single reviewer and the remainder of the items were scored by two independent reviewers. Any disagreements between reviewers were resolved by discussion.

Data analysis

The proportion and 95% confidence interval of included studies that were rated as 'yes' on each TIDieR item were calculated for the intervention and control groups (separately). The proportions were also used to present the relationship between the number of TIDieR items met for each study and the methodological quality (PEDro score) and by subdisciplines coded on PEDro. The relationship between the completeness of interventions and methodological quality is presented using bubble graphs, where the size of each bubble represents the proportion of trials that share the same combination of TIDieR and PEDro scores, for both intervention and control groups. We also calculated the proportion of studies meeting each TIDieR item by subdiscipline for both intervention and control groups.

The Prevalence and Bias Adjusted Kappa (PABAK) coefficient was used to quantify the reliability for each item of the TIDieR checklist, with separate analyses for the intervention and control groups. Interpretation of the PABAK coefficient has been described as: <0 = poor, 0.00 to 0.20 = slight, 0.21 to 0.40 = fair, 0.41 to 0.60 = moderate, 0.61 to 0.80 = substantial, 0.81 to 1.00 = almost perfect [16]. All data were analysed using Microsoft Office Excel 2010.

Results

The included studies were sampled from the 4 August 2014 version of PEDro. In this version, PEDro contained 28 042 reports of randomised controlled trials, systematic reviews and evidence-based clinical practice guidelines. Of the 22 617 trial reports, 1492 were published in English in

Table 1
Prevalence and Bias Adjusted Kappa values for each TIDieR item for the intervention and control groups for the two sub-samples of included studies used to evaluate inter-rater reliability.

		First sub-sample		Second sub-sample	
TIDieR item		**Intervention**	**Control**	**Intervention**	**Control**
1. Intervention name		0.82	1.00	1.00	1.00
2. Intervention rationale		0.88	0.52	0.93	0.84
3. (materials)	**3a.** Intervention materials	0.82	0.58	0.47	0.60
	3b. Location of materials	0.76	0.15	0.53	0.76
4. Intervention procedures		0.70	0.58	0.47	0.92
5. Who provided intervention		0.03	0.15	0.93	0.76
6. (how)	**6a.** Mode of delivery	0.94	0.58	0.87	0.76
	6b. Individual or group	0.52	0.27	0.80	0.68
7. Intervention location		0.39	0.52	0.60	0.52
8. (when and how much)	**8a.** N° of sessions	0.36	0.15	0.67	0.76
	8b. Schedule of intervention	0.70	0.27	0.87	0.76
	8c. Duration of intervention	0.52	0.27	0.73	0.92
	8d. Intensity or dose	0.58	0.45	0.33	0.60
9. Tailoring of intervention		0.52	0.70	0.60	0.84
10. Modification of intervention during trial		1.00	0.70	1.00	1.00
11. How the intervention fidelity assessed		0.76	0.58	0.53	0.76
12. Actual intervention fidelity		0.58	0.70	0.53	0.92

Items with 'substantial' or 'almost perfect' reliability are shaded in grey.

2013 and had total PEDro score and subdiscipline coding. These 1492 trial reports were exported to Microsoft Excel and were sorted in random order using a randomisation formula; then the first 200 trial reports from this list that fulfilled the inclusion criteria were included in this study. The average total PEDro score for the 200 included studies was 5.6 points (±1.51). Most studies included were classified within the subdisciplines of gerontology (24%), musculoskeletal (20%), and continence and women's health (17%). The subdisciplines of ergonomics, orthopaedics, and sports represented less than 5% of the studies. The remaining areas accounted for 9 to 15% of the included studies.

Inter-rater reliability for each item in the TIDieR checklist (separately for the intervention and control arms) was evaluated in two sub-samples of 30 included studies. The PABAK values for the first sub-sample (studies 7 to 36 – studies 1 to 6 we did the pilot) were classified as 'substantial' or 'almost perfect' (i.e. 0.61 or higher) for just 13 of the 34 TIDieR items (see Table 1). Inter-rater reliability improved for the second sub-sample (studies 67 to 96 – studies 37 to 66 were discussed and rated to improve consensus), with PABAK values ≥0.61 for 23 of the 34 TIDieR items (see Table 1). Items with PABAK values ≤0.60 in the second sub-sample were rated independently by two reviewers (with other items assessed by a single reviewer) for studies 97 to 200.

The completeness of intervention reporting was higher for the intervention conditions (77% of reports fulfilling at least half of the TIDieR items) than for the control conditions (26%). The proportions and 95% confidence intervals for achieving a 'yes' for each TIDieR item are presented in Table 2. For the intervention conditions, the items for which >75% were rated as 'yes' were: the intervention name (item 1), intervention rationale (item 2), mode of delivery (item 6a), and details of the 'when and how much' of the interventions (items 8a to 8c). The items that rated the poorest (<25% rated as 'yes') were: information about access to intervention materials (item 3b), any intervention modifications (item 10), and assessment or monitoring of intervention fidelity (item 11).

For the control conditions, only one item was rated as 'yes' in >75% trials: the intervention name (item 1). The items with the poorest descriptions ('yes' in <25% of trials) were: intervention rationale (item 2), details of intervention materials and where to access them (items 3a and 3b), intervention dose or intensity (item 8d), any planned tailoring (item 9) or unplanned modifications (item 10) of the control intervention, and if the fidelity of the control intervention was assessed, monitored or delivered as planned (items 11 and 12). In addition, 46 studies were rated as 'not applicable' for control groups, as there was no intervention provided.

Fig. 1 shows the proportion of the TIDieR items met by the trials in each subdiscipline of physiotherapy (NB some trials were coded as relevant to more than one subdiscipline so trial count is >200). The description of the treatments for the intervention group achieved a 'yes' for 50% to 65% of all TIDieR items, with the exception of the sports

Table 2
Number, proportion (95% confidence interval) of achieving a 'yes' for each TIDieR item in both intervention and control groups.

Items		Intervention group *n*, % (95% CI)	Control group *n*, % (95% CI)
1. Intervention name		200, 100 (98 to 100)	198, 99 (96 to 100)
2. Intervention rationale		190, 95 (91 to 97)	37, 19 (14 to 25)[a]
3. (Materials)	3a. Intervention materials	80, 40 (34 to 47)	47, 24 (18 to 30)[a]
	3b. Location of materials	41, 21 (16 to 27)	19, 10 (6 to 14)[a]
4. Intervention procedures		117, 59 (52 to 65)	56, 28 (22 to 35)[a]
5. Who provided intervention		108, 54 (47 to 61)	61, 31 (25 to 37)[a]
6. (How)	6a. Mode of delivery	197, 99 (96 to 100)	107, 54 (47 to 60)[a]
	6b. Individual or group	124, 62 (55 to 68)	63, 32 (26 to 38)[a]
7. Intervention location		129, 65 (58 to 71)	88, 44 (37 to 51)[a]
8. (When and how much)	8a. No. of sessions	177, 89 (83 to 92)	95, 48 (41 to 54)[a]
	8b. Schedule of intervention	183, 92 (87 to 95)	95, 48 (41 to 54)[a]
	8c. Duration of intervention	160, 80 (74 to 85)	62, 31 (25 to 38)[a]
	8d. Intensity or dose	108, 54 (47 to 61)	44, 22 (17 to 28)[a]
9. Tailoring of intervention		66, 33 (27 to 40)	6, 3 (1 to 6)[a]
10. Modification of intervention during trial		0, 0 (0 to 2)	0, 0 (0 to 2)[a]
11. How intervention fidelity was assessed		22, 11 (7 to 16)	12, 6 (4 to 10)[a]
12. Actual intervention fidelity		87, 44 (37 to 50)	31, 16 (11 to 21)[a]

[a] Item rated as 'not applicable' to 46 (23%; 95% CI 18 to 29) of the included study control groups.

subdiscipline. In contrast, the descriptions of the treatments for the control groups were less complete for all subdisciplines, with between 30% and 40% of TIDieR items being achieved for only six subdisciplines (sport, orthopaedics, oncology, neurology, musculoskeletal and continence and women's health).

The association between completeness of reporting of trial interventions (measured using TIDieR) and trial

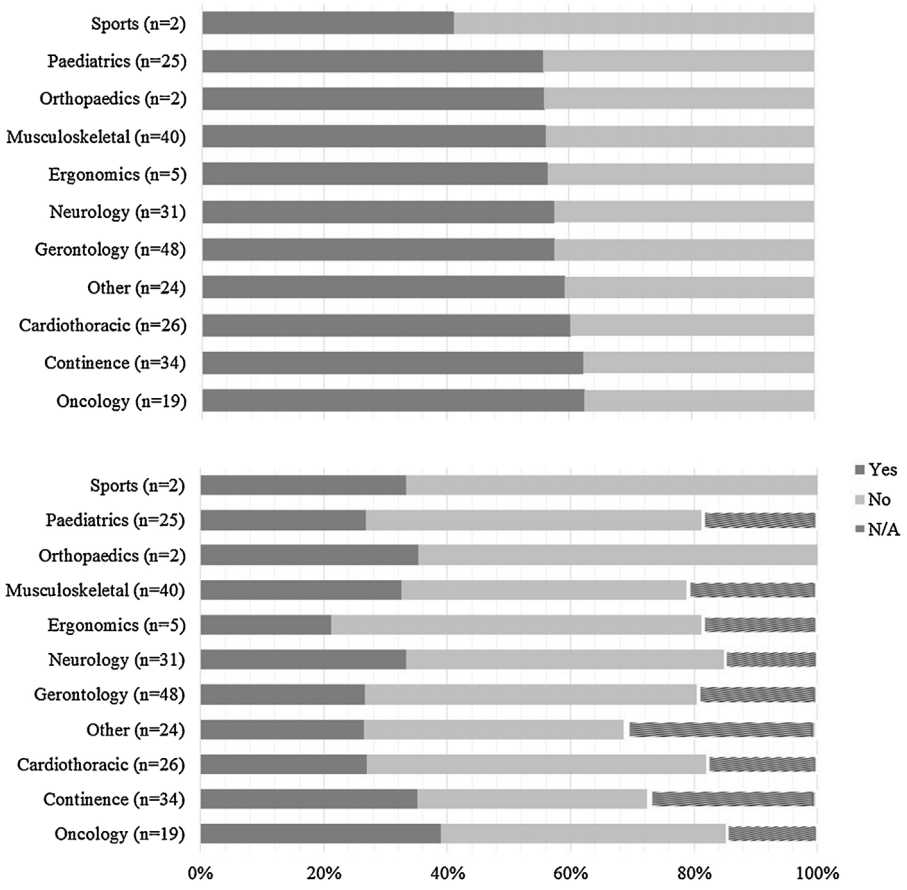

Fig. 1. Percentage of TIDieR items achieving a 'yes', 'no' or 'not applicable' (N/A) rating for each subdiscipline of physiotherapy. Top panel is the intervention groups. Bottom panel is the control groups.

quality (measured using the PEDro scale) is presented in bubble plots (Fig. S2). Visual inspection of the plots suggests that there is no clear relationship between methodological quality and completeness of the description of the intervention for the intervention or control groups.

Discussion

The completeness of reporting of interventions in trials evaluating physiotherapy interventions was generally poor. For the experimental intervention conditions, 46 (23%) trials did not describe at least half of the TIDieR items. Reporting was worse for the control conditions, of which 149 (75%) trials described less than half of the items. The items that were most frequently poorly described were: the materials used in the intervention, the amount of intervention provided (number of sessions/schedule/dose/intensity), tailoring/modification of the intervention and intervention fidelity. There was no clear difference in the completeness of descriptions of interventions across the physiotherapy subdisciplines, and no obvious relationship between the completeness of intervention reporting and trial methodological quality.

This is the first study evaluating the descriptions of interventions in a sample of physiotherapy trials. The strengths of this study include a large and representative sample of trials and the use of two independent reviewers to evaluate the TIDieR items that had low reliability. A limitation of the study was that we only included trials published in English in 2013, so our findings may not generalise to studies in other languages or earlier studies. This study did not contact authors and explore why the interventions were poorly described in physiotherapy trial reports.

The problem of incomplete reporting of interventions that we observed in physiotherapy trial reports also occurs in other areas of healthcare. A recent study reported the inadequate description of interventions in over 60% of trials of non-pharmacological interventions which included surgery, technical procedures, devices, rehabilitation, psychotherapy, behavioural interventions, and complementary and alternative medicine [3]. These findings are consistent with our study and with previous studies reporting significant gaps in intervention reporting that would compromise a clinician's ability to replicate the intervention for their patients [5–8].

A particular problem in the physiotherapy discipline is that many interventions are described by the name of the developer of the intervention (e.g. Pilates, McKenzie therapy, Alexander Technique) or a presumed mechanism of action (e.g. stabilisation exercise). The issue is that there are no agreed standards for the use of these treatment names and the same term can be applied to quite different intervention protocols.

Describing the control intervention as 'usual care' or 'conventional' treatment obscures accurate interpretation of the effect size achieved by the intervention. These control interventions can vary enormously depending on factors such as the country, care provider, or clinic service routine. These issues with reporting also hinder synthesis and comparisons between studies in systematic reviews and evidence-based clinical practice guidelines.

To address the problem of poor reporting of interventions in physiotherapy trial reports we recommend that the TIDieR checklist be adopted by journals as a mandatory tool to guide authors and reviewers, in parallel with other statements such as CONSORT [10], the CONSORT extension for non-pharmacological intervention [17] and SPIRIT [11]. We recognise that many journals have word limits; however, electronic or online resources can be used to provide a more complete description of the interventions [18]. For more complex interventions that are common in physiotherapy, the use of videos or websites demonstrating the interventions used can also be considered. Nevertheless, authors should ensure that these supplementary materials are cited in the trial reports and are accessible (e.g. freely accessible, URL working). For example, Sakakibara *et al.* 2013 [19] developed a website to report the wheelchair skills training program used in their trial, and Janyacharoen *et al.* 2013 [20] included a supplementary video of the intervention. These initiatives will reduce waste in research and improve the usability of future trials.

Inadequate reporting of interventions is a problem in many areas of health care. There is a need for strategies to improve value and reduce waste in research studies and better reporting would contribute to reduced waste by enhancing the reproducibility of interventions [21,22]. The responsibility for improving the quality of reporting should not be seen as the sole responsibility of authors, with journal editors, peer reviewers, and funding agencies all able to contribute solutions to this major problem. TIDieR can guide research authors, editors, reviewers, and clinicians on how to better report, assess, and understand the interventions in different stages of publication, including protocols and at peer review. We also argue that the assessment of the impact of a scientist's research should be expanded beyond traditional metrics such as number of publications and citations and also consider the completeness of reporting to allow application of their research. We need good quality research that can also be used appropriately. Future research could evaluate the effectiveness of strategies to optimise the completeness of descriptions of interventions such as the hindrance of the publication unless the intervention is fully described.

Acknowledgements

Tiê Parma Yamato is supported by CAPES (Coordination for the Improvement of Higher Education Personnel), Brazil. Bruno Tirotti Saragiotto is supported by CNPQ (National Council for Scientific and Technological

Development), Brazil. Chris G. Maher is supported by a Principal Research Fellowship from the National Health and Medical Research Council, Australia. Tammy C. Hoffmann is supported by a National Health and Medical Research Council of Australia Career Development Fellowship.

Ethical approval: None required.

Conflict of interest: None declared.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.physio.2016.03.001.

References

[1] Schroter S, Glasziou P, Heneghan C. Quality of descriptions of treatments: a review of published randomised controlled trials. BMJ Open 2012;2:e001978.
[2] Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, *et al.* Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655.
[3] Hoffmann TC, Erueti C, Glasziou PP. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. BMJ 2013;347:f3755.
[4] Herbert RD, Bo K. Analysis of quality of interventions in systematic reviews. BMJ 2005;331:507–9.
[5] Duff JM, Leather H, Walden EO, LaPlant KD, George Jr TJ. Adequacy of published oncology randomized controlled trials to provide therapeutic details needed for clinical application. J Natl Cancer Inst 2010;102:702–5.
[6] Jacquier I, Boutron I, Moher D, Roy C, Ravaud P. The reporting of randomized clinical trials using a surgical intervention is in need of immediate improvement: a systematic review. Ann Surg 2006;244:677–83.
[7] Robb SL, Carpenter JS. A review of music-based intervention reporting in pediatrics. J Health Psychol 2009;14:490–501.
[8] Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, *et al.* Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015;350:h1258.
[9] Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, *et al.* Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687.
[10] Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.
[11] Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleza-Jeric K, *et al.* SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013;158:200–7.
[12] Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med 2010;7:e1000217.
[13] Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther 2003;83:713–21.
[14] Macedo LG, Elkins MR, Maher CG, Moseley AM, Herbert RD, Sherrington C. There was evidence of convergent and construct validity of Physiotherapy Evidence Database quality scale for physiotherapy trials. J Clin Epidemiol 2010;63:920–5.
[15] de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. Aust J Physiother 2009;55:129–33.
[16] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
[17] Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, Group C. Methods and processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments. Ann Intern Med 2008;148:W60–6.
[18] Hoffmann T, English T, Glasziou P. Reporting of interventions in randomised trials: an audit of journal instructions to authors. Trials 2014;15:20.
[19] Sakakibara BM, Miller WC, Souza M, Nikolova V, Best KL. Wheelchair skills training to improve confidence with using a manual wheelchair among older adults: a pilot study. Arch Phys Med Rehabil 2013;94:1031–7.
[20] Janyacharoen T, Laophosri M, Kanpittaya J, Auvichayapat P, Sawanyawisuth K. Physical performance in recently aged adults after 6 weeks traditional Thai dance: a randomized controlled trial. Clin Interv Aging 2013;8:855–9.
[21] Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, *et al.* Increasing value and reducing waste in research design, conduct, and analysis. Lancet 2014;383:166–75.
[22] Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, *et al.* Reducing waste from incomplete or unusable reports of biomedical research. Lancet 2014;383:267–76.

Available online at www.sciencedirect.com

ScienceDirect