Journal of Physiotherapy 69 (2023) 65–67

AUSTRALIAN PHYSIOTHERAPY ASSOCIATION

Journal of PHYSIOTHERAPY

journal homepage: www.elsevier.com/locate/jphys

Editorial

Are we giving stratified care a fair trial?

Nadine E Foster [a,b], Jonathan C Hill [b], Jesper Knoop [c]

[a]STARS Education and Research Alliance, Surgical Treatment and Rehabilitation Service (STARS), The University of Queensland and Metro North Health, Brisbane, Australia; [b]School of Medicine, Keele University, Keele, UK; [c]Department of Health Sciences, Amsterdam Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

The recent OCTOPuS study in the *Journal of Physiotherapy* is unique, being the first randomised controlled trial (RCT) to test stratified exercise therapy for patients with knee osteoarthritis.[1] Stratified care involves targeting treatment to subgroups of patients based on key characteristics such as prognostic factors, responsiveness to treatment and/or underlying mechanisms.[2] Most stratified care trials to date have focused on people with low back pain – such as the UK STarT Back trial,[3] the MATCH[4] and TARGET[5] trials in the USA, and trials from Denmark[6] and the Netherlands[7] – but OCTOPuS is the first stratified care trial to have tested a model of stratified care in people with knee osteoarthritis. Almost two decades ago stratified care was called the 'Holy Grail' of back pain research,[8] and stratified care or precision medicine has become the zeitgeist in many clinical fields – such as mental health,[9] diabetes[10] and cancer[11] – with the goal of personalised healthcare. The allure is the potential to maximise treatment benefit, reduce harm and increase healthcare efficiency by offering the right treatment to the right patient at the right time.[12] However, the findings of the OCTOPuS trial are not unique, as they concluded that stratified care did not lead to better clinical[1] or cost outcomes.[13]

This Editorial considers one particularly compelling explanation for the lack of differences observed in stratified care trials: intervention fidelity. It highlights the importance of the two key components of stratified care (subgrouping and targeted treatment) and makes recommendations for future research to determine feasibility and optimise intervention fidelity before conducting full-scale RCTs.

The stratified care model in OCTOPuS was based on empirically observed phenotypes (distinct clinical profiles), a valid stratification algorithm (robust categorisation of participants into subgroups) and previously effective exercise interventions (evidence-based targeted treatment options).[1] Participants were phenotyped into three subgroups for targeted treatment: the 'high muscle strength' subgroup was matched to a targeted program of three to five home-based exercise sessions over 12 weeks and one booster session; the 'low muscle strength' subgroup was matched to eight to 12 supervised sessions plus home exercise sessions and one to two booster sessions; and the 'obesity subgroup' was matched to 12 to 18 individual exercise sessions with two to three booster sessions plus five to eight dietitian sessions and at least one joint consultation with both a physiotherapist and dietitian. The trial compared a stratified approach versus usual exercise therapy involving an average of 10 physiotherapy sessions of education and exercises designed around strength, fitness, function and balance/stabilisation training. The OCTOPuS results showed no added value of stratified care for clinical[1] or cost-effectiveness outcomes.[13] Frequently reported protocol violations in the stratified care arm were the delivery of too few physiotherapy and dietary sessions, no booster sessions and barriers to interprofessional collaboration.[1,14]

Key questions to ask of trials of stratified care are 'Were the participants successfully subgrouped?' and 'Did they receive the treatment matched to their subgroup?' Imagine a cancer trial, for example, that wanted to test whether people with a certain clinical profile (subgroup) have better outcomes if they receive a specific matched medication (targeted treatment). If trial participants were not subgrouped in the intended way or did not receive the matched medication, the trial would not be a fair test of stratified care. Similarly, in physiotherapy, where upskilling or training of healthcare providers is needed for both components of stratified care (subgrouping and targeting treatment), trials can fail to be a 'fair test' of stratified care. Below we consider these issues using previous trials of stratified care.

The STarT Back risk stratification trial for low back pain from the UK demonstrated positive results in favour of stratified care.[3] All participants were successfully subgrouped (using the STarT Back tool) and only 2.6% of those randomised to stratified care did not receive the matched treatment for their subgroup. In addition, given that different therapists delivered the control and intervention treatments, contamination between arms of the trial was avoided. To achieve high intervention fidelity in the STarT Back trial, dedicated new clinics using trained research staff were set up, the stratified care treatment (three subgroups each matched to one treatment protocol) was provided by a small team of 13 trained physiotherapists, who were given time for self-reflection and skills practice as part of their training, and were supported by ongoing mentoring, feedback, and jointly reviewing the management plans of example patients.[15,16] These processes had all been informed by a pilot study before the main STarT Back trial that demonstrated the feasibility and acceptability of the approach.[17]

Intervention fidelity has been a challenge in all subsequent stratified care trials in the field. A subsequent quality improvement study in the UK (the IMPaCT Back study) showed similar results as the STarT Back trial but in real-life general practice,[18] although the positive results were attenuated by the lower intervention fidelity observed (general practitioners followed the risk stratification tool's recommendation for matched treatment in 71% of participants). The MATCH trial in the USA tested an adapted STarT Back approach but, despite education and training of clinicians and embedding the subgrouping tool into the electronic health record, only about half of participants were successfully subgrouped and there was no evidence that clinical management changed based on the subgroup.[4] The TARGET trial in the USA focused on acute, high-risk patients with low back pain (ie, one patient subgroup) for whom the intended matched intervention was psychologically informed physiotherapy.[5] Despite creating an automated process to identify the subgroup and generate psychologically informed physiotherapy referrals, only 36% of participants in the stratified care arm were referred for psychologically informed physiotherapy and 30% of controls were also referred for it. Intervention fidelity was well below expectations, leading the TARGET trial authors to reflect that they had 'balanced the

https://doi.org/10.1016/j.jphys.2023.02.014

pragmatic-explanatory continuum of their RCT more towards pragmatic with regard to flexibility in delivery of the intervention'.[5]

A Danish trial of stratified care amended the STarT Back approach for low back pain for their setting, with training for healthcare providers supported by the same training lead from the original UK STarT Back trial.[6] That trial found no positive effects from stratified care but noted that treatment was similar in the two trial arms and that, due to slow recruitment, clinicians reported insufficient opportunities to practice and develop their confidence and competence in delivering the matched treatments.[6]

The UK STarT MSK trial compared an adapted model of stratified versus usual primary care for adults with the five most common musculoskeletal pain presentations.[19] General practitioners were introduced to stratified care through training sessions, facilitated by embedding the STarT MSK subgrouping tool in electronic pop-up computer templates that recommended matched treatment options for each participant (there were 15 matched treatment options), and general practitioners were invited to share and discuss feedback on their intervention fidelity. Whilst some aspects of clinical decision-making about treatment changed in a positive direction, there were no overall benefits for participants' clinical outcomes. The general practitioners used the subgrouping tool for only 30% of all eligible participants, highlighting the challenge of incorporating risk stratification into their workflows. Only those who completed questionnaires, in which the STarT MSK tool was embedded, and provided consent to participate were included in the trial, and in the stratified care arm general practitioners reported selecting a matched treatment for 77% of those participants.[19] Whilst treatment fidelity was moderate, it is unlikely that this alone was the reason for the lack of positive findings. Other explanations could be the lack of the effectiveness of the matched treatments, and the way in which the stratification was restricted to the initial primary care decision alone, without any further system-level changes in the services that participants were referred to, given their subgroup.

Lastly, a Dutch stratified care trial found conflicting results with positive reductions for pain intensity only.[7] The trial compared a new stratified blended intervention for low back pain (stratified using the STarT Back tool) comprising a smartphone app integrated with face-to-face physiotherapy versus face-to-face physiotherapy alone. The authors noted that there were few differences in the treatment across the two trial arms, again suggesting that treatment targeting (where treatment is different in stratified care than control and is different for different subgroups of patients) was not successfully achieved.[7]

With the addition of the OCTOPuS stratified care trial,[1] there is a pattern emerging of similar conclusions that stratified care is not clinically effective. Is it time to give up on stratified care or are we failing to give it a fair trial? Perhaps stratified care is truly not superior to usual care. As proposed by Knoop et al,[1] usual physiotherapy may not be a 'one-size-fits-all' approach but already personalised, for example in ways that better harness the understanding of the importance of an individual patient's beliefs, so that every patient receives a treatment targeting their personal needs, obstacles to recovery and preferences. Thus, no further models of stratified care are needed. However, if future research aims to develop and test new models of stratified care that better target treatment to subgroups of participants, the field should learn from this suite of trials. Stratified care requires changes to clinical workflows, training of healthcare providers and system-level changes such as different referral processes. Yet, to date, stratified care trials have: lacked monitoring or feedback to encourage clinicians to deliver matched treatments with high fidelity;[1,20] included numerous treatment options, some of which are difficult to access;[4,5,6,20] experienced contamination between intervention arms;[1,5,6] and lacked contrast in the matched treatments for different subgroups.[6,19]

It could be argued that if stratified care interventions are so complex to deliver in the real world, even if a carefully controlled explanatory RCT shows benefit, then the intervention is likely not worthwhile anyway. However, this is rather like the conundrum of clinical guidelines. Guidelines review individual RCTs (or reviews of individual RCTs), offer recommendations for best clinical practice based on the results of those RCTs that were, for the most part, explanatory trials. Yet most studies that evaluate the effect of implementation strategies of guideline recommendations in RCTs show no benefit on patient clinical outcomes.[21–23] We don't tend to then conclude that guidelines are unhelpful, but that we have yet to learn much about the art and science of successful implementation into complex healthcare systems.[24] Another learning might be that for trials of stratified care (or indeed any complex intervention), the sample size should be inflated for the likely lower rates of intervention fidelity that are typically seen in RCTs conducted in real-world settings. However, this may make conducting such trials neither feasible nor fundable.

It is therefore proposed that future attempts to test the effectiveness of stratified care need to give stratified care a 'fair trial'. It is suggested that more careful staging of the research is needed (rather like in the cancer field), whereby efficacy studies that are more highly controlled are needed first (to ensure that participants with the phenotype (subgroup) are clearly identified and then receive the treatment they are supposed to), to determine whether there is merit in the intervention and, only if so, then subsequent studies focus on the feasibility of delivering the stratified care intervention. Only when feasibility, at least initially, is shown or improved such that there is adequate intervention fidelity, should larger and more expensive RCTs be conducted that test the clinical effectiveness of the stratified care intervention on clinical outcomes with attention to careful intervention fidelity, followed by studies that then focus on implementation in real-world clinical settings as the primary aim (such as hybrid type 3 implementation trials).[24] Such a staged approach is similar to the phases recommended in the framework for the development and evaluation of complex interventions.[25] New trial designs, such as adaptive platform trials, might mean that it becomes more possible to add in new subgroups with matched treatment arms without stopping and starting a new trial (with all the expense and effort)[26] but intervention fidelity will continue to be a pervasive challenge. It is recommended that researchers, clinicians and funders better understand and mitigate the challenge of intervention fidelity in future RCTs of stratified care. Tools such as the PRagmatic Explanatory Continuum Indicator Summary-2 (PRECIS-2) Provider Strategies could be helpful aids in the design of future stratified care trials.[27]

Footnotes: Nil.

Ethics approval: Not applicable.

Competing interests: The authors are active researchers in the field of stratified care.

Source(s) of support: Nil.

Acknowledgements: Nil.

Provenance: Not invited. Peer reviewed.

Correspondence: Nadine E Foster, STARS Education and Research Alliance, Surgical Treatment and Rehabilitation Service (STARS), The University of Queensland and Metro North Health, Brisbane, Australia. Email: n.foster@uq.edu.au

References

1. Knoop J, et al. *J Physiother*. 2022;68:182–190.
2. Foster NE, et al. *Best Prac Res Clin Rheum*. 2013;27:649–661.
3. Hill JC, et al. *Lancet*. 2011;378:1560–1571.
4. Cherkin D, et al. *J Gen Intern Med*. 2018;33:1324–1336.
5. Delitto A, et al. *E Clin Med*. 2021;34:100795.
6. Morsø L, et al. *Eur J Pain*. 2021;25:2020–2038.
7. Koppenaal T, et al. *J Med Internet Res*. 2022;24:e31675.
8. Waddell G. *J Rheum*. 2005;32:395–396.
9. Fletcher S, et al. *Lancet Psychiatry*. 2021;8:202–214.
10. Brandon R, et al. *Front Endocrinol*. 2023;13:1091421.
11. Keller RB, et al. *JCO Precis Oncol*. 2023;7:e2200342.
12. Hingorani AD, et al. *BMJ*. 2013;346:e5793.
13. Knoop J, et al. *Braz J Phys Ther*. 2022;27:100469.
14. Knoop J, et al. *BMC Musculoskelet Disord*. 2022;23:559.
15. Main CJ, et al. *Physiotherapy*. 2012;98:110–116.
16. Sowden G, et al. *Fam Pract*. 2012;29:50–62.
17. Hay EM, et al. *BMC Musculoskelet Disord*. 2008;9:58.
18. Foster NE, et al. *Ann Fam Med*. 2014;12:102–111.
19. Hill JC, et al. *Lancet Rheum*. 2022;4:e591–e602.

20. Hsu C, et al. *Pain Med*. 2019;20:1105–1119.
21. Fillipo R, et al. *Implement Sci Commun*. 2022;3:57.
22. Suman A, et al. *Implement Sci*. 2016;11:126.
23. Dziedzic KS, et al. *Osteo Cartil*. 2018;26:43–53.
24. Hassett L, et al. *J Physiother*. 2022;3:210–212.
25. Skivington K, et al. *BMJ*. 2021;374:n2061.
26. Pallmann P, et al. *BMC Med*. 2018;16:29.
27. Norton WE, et al. *Implement Sci*. 2021;16:7.