BRUNO T. SARAGIOTTO, MSc[1] • CHRIS G. MAHER, PT, PhD, FACP[1] • MARK J. HANCOCK, PhD[2] • BART W. KOES, PhD[3]

Subgrouping Patients With Nonspecific Low Back Pain: Hope or Hype?

J Orthop Sports Phys Ther 2017;47(2):44-48. doi:10.2519/jospt.2017.0602

Clinicians and clinical researchers share a common goal of achieving better outcomes for patients with low back pain (LBP). For that, randomized controlled trials and systematic reviews are the most reliable study designs to determine the effects of interventions. Subgroup analyses in these research designs have been used to examine treatment-effect modification across subgroups defined by patient characteristics. The goal of this approach is to identify patient-level factors associated with greater effects of treatment than those which occur on average. Armed with this information, clinicians can match individual patients to the treatment that will be most effective for them.

The identification of patient subgroups that respond best to specific interventions has been set as a key priority in LBP research for the past 2 decades.[2,7] In parallel, surveys of clinicians managing LBP show that there are strong views against generic treatment and an expectation that treatment should be individualized to the patient.[6,22] However, despite this emphasis on treatment-based subgroups, little high-quality evidence exists for the investigation of subgroups of patients with LBP who respond best to specific interventions.[31,34]

Calls for caution when reading reports involving subgroups have been made repeatedly in the general medical literature over the last 30 years, but these papers do not seem to have dulled the enthusiasm in the LBP field. We feel that it is timely to consider whether a prolonged focus on identifying subgroups is useful or not. To meet this challenge, we will consider the advantages and potential problems surrounding subgroup analyses in nonspecific LBP. This will allow more informed decisions on whether more or less emphasis, time, and resources should be dedicated to identifying treatment subgroups for LBP.

In this article, we present supporting and opposing arguments for the subgrouping approach in nonspecific LBP, considering the progress made so far in the LBP field and the relevant literature in adjacent fields. We have deliberately chosen to argue 2 extreme positions (pros and cons), as is done with an Oxford debate, so readers can make better conclusions considering the advantages and disadvantages of subgrouping.

Viewpoint: The Investigation of Subgroups Is Important and Useful to Advancing Knowledge and Clinical Practice in LBP

One Size Fits All Does Not Work Well Nonspecific LBP accounts for the great majority of cases of LBP and is defined as LBP for which there is no identifiable cause (eg, injury or disease).[1] As a result, treatment recommendations commonly involve a one-size-fits-all approach.[25] Following this approach in clinical trials, most treatment options tested provide small or even no benefit.[1,36]

While the optimal treatment for nonspecific LBP is still unknown, the burden of the condition remains massive.[17] The current treatment classification system (ie, a small group [5%-10%] of patients with identified specific pathology versus the large group [90%-95%] with nonspecific LBP) is clearly not working well. The large group of nonspecific LBP patients

[1]Musculoskeletal Division, The George Institute for Global Health, Sydney Medical School, The University of Sydney, Sydney, Australia. [2]Faculty of Medicine and Health Sciences, Macquarie University, Sydney, Australia. [3]Department of General Practice, Erasmus University Medical Center, Rotterdam, the Netherlands. The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Dr Chris G. Maher, The George Institute for Global Health, PO Box M201, Missenden Road, NSW 2050 Australia. E-mail: cmaher@georgeinstitute.org.au

does not seem to be homogeneous; instead, it probably includes patients with quite different underlying contributors to their condition.[5,22] Likewise, patients with nonspecific LBP show marked differences in the course of their back pain.[14] These individual variations and needs should be taken into account to develop a tailored treatment program, which is likely to produce optimal outcomes for each individual. The investigation of subgroups of people with different features who would respond better to one treatment than to another offers the possibility of larger treatment effects.[8,15,16,21]

Methods Are Improving The research methods for validly investigating specific treatment effects for subgroups of patients have been well developed in the past decades. Published guidelines and overviews for conducting subgroup analyses in randomized trials[10,24,35] indicate that methods are indeed improving in this area. The **TABLE** shows the key methodological issues for validly investigating treatment effects in subgroups. It is true that the valid methods for investigating subgroups have not been applied widely in available studies and research papers. There are, however, also examples of well-planned and well-executed initiatives.[13,19] The current lack of validation studies is a problem for the implementation of treatment subgroups but lays the foundation for better high-quality prospective validation studies in the future.[33] Hypothesis-setting studies of lower quality are important to guide the development of large high-quality studies and are not themselves a problem (or low quality) as long as they are correctly and cautiously interpreted.

It Does Not Need to Be Complex Subgrouping patients in LBP does not need to be complex or difficult. Sometimes in medicine, the most powerful subgroups are based on one feature (eg, management of a stroke based on whether the stroke is due to a bleed or a clot). A good example in the LBP field is the STarT Back trial that used a simple prognostic tool (9 questions only) to match patients to treatment packages appropriate for them.[19] Importantly, simple stratified primary care management showed better clinical and economic results than usual care. This finding is encouraging. Simple, easy-to-apply tools for subgrouping and managing patients may well facilitate implementation in clinical practice.

Fits the Personalized Medicine Approach Personalized medicine is a "new" approach to individualize interventions for prevention and treatment. This is based on the classification of patients into different subpopulations depending on their characteristics (ie, genetic markers, susceptibility to disease, or response to treatment), with the aim to optimize efficacy.[3] In the past, particularly the field of genetics has championed this new approach, in which genetic and molecular information is used to determine predisposition to a particular health condition, to confirm a diagnosis, or to individualize treatment (ie, the right treatment for the right patient at the right time).[29] This approach has been reported to be effective for some areas in medicine (eg, use of trastuzumab based on the HER2/neu test result to treat breast cancer).[32]

Accordingly, individual (personalized) treatment may well be based on demographic and clinical characteristics measured during the history taking or physical examination, or from laboratory or radiological tests (eg, imaging features). Subgrouping patients with nonspecific LBP fits very well into this approach, and over time we may include gene variation in our subgroups as well.

The Subgroup Approach Is Preferred by Clinicians (and Patients) Clinicians are usually favorable to the idea of individualized treatments for nonspecific LBP. It is common practice for clinicians to use subgroup labels for describing nonspecific LBP patients.[23] Clinicians assert that they recognize in their practice various subgroups of patients with a similar clinical course and reaction to therapy. When asked, 93% of clinicians do not think nonspecific LBP is one condition, and about 75% think that it is possible to recognize LBP subgroups.[22] Additionally, the broad term *nonspecific* does not guide treatment choices very well. It has thus been reported to not be useful and to be of limited relevance to practice by clinicians.[6]

At a more basic level, one of the principles of evidence-based practice is that clinicians should use their day-to-day knowledge (ie, clinical experience) in clinical decision making. Clinical experience may suggest specific approaches for patients with particular presentations. When the great majority of clinicians feel that patients with nonspecific LBP are not one group, they are compelled to try to use their experience to individualize intervention until there is good evidence to guide this. Regarding patient preferences, another principle of evidence-based practice, it seems that patients recognize their individual variability in response to treatments when indicating a preferred

TABLE Key Methodological Features for Investigating Treatment Effects in Subgroups[31,34]

- The subgroup variable should be a characteristic measured at baseline.
- A subgroup analysis must be preplanned to test a hypothesis, and it should be specified a priori (ie, study protocol, primary trial).
- The subgroup analysis should be carried out based on a small number of hypotheses tested (preferably fewer than 5).
- Statistical tests of significance should be used to assess the likelihood that a given interaction might have arisen due to chance alone. If multiple significant interactions exist, this needs to be tested for independence.
- The subgroup effect should be consistent with evidence from previous studies (ie, replication).
- The subgroup effect should be consistent across related outcomes.
- There should be a strong pre-existing biological rationale supporting the apparent subgroup effect (or indirect evidence).

treatment. Patients commonly report cases in which one treatment worked for them but another did not.

Counterpoint: Subgroup Analysis Is Misguided and a Misleading Distraction From Improved Understanding of LBP

Poor Methodological Quality Subgroup analyses evaluating differential response to treatments for a range of health conditions have been reported to commonly have poor methodological quality and low credibility.[34] Although identification of treatment subgroups has been set as a key research priority in the LBP field for the last 15 years,[2,7] study quality has remained low over this period.[18,31] Most subgroup analyses still do not specify the hypothesis a priori and do not conduct interaction tests (treatment by effect modifier).[31,34] Lack of power in subgroup analyses is another problem; it is always lower than the power for the main effect (if power for the main effect is 80% to 90%, for a subgroup it will always be less, about 20% to 30%).[10,26] In addition, many studies include multiple comparisons, increasing the chance of a false-positive result.[10,34] Besides the problems with analyses, there is a common reporting problem in which authors commonly overstate their claims of subgroup effect beyond what the study design and results would justify.[31]

Biological Rationale Is a Challenge for Treatment of Nonspecific LBP It is accepted that findings of treatment-effect modification in a trial should be considered more credible if the result has a biological rationale.[28,35] That is, we should place greater faith in treatment-effect modification that is consistent with our current understanding of the biologic mechanisms of disease and the mechanisms of treatments. In the case of nonspecific LBP, this assumption is difficult to fulfill because there is no identified biological source of nonspecific LBP, and many contemporary treatments have an unclear mechanism of action (eg, spinal manipulative therapy, exercise). In the breast cancer area, for example, the recent discovery of specific prognostic and predictive biomarkers that enable the application of more individualized therapies for these patients has changed treatment over the past decades. This seems to be possible when the biological source is known and the treatment action understood. As we know so little about causes of LBP and the mechanisms of treatments, it is difficult to have any strong rationale for treatment-effect modification.

Qualitative Heterogeneity of Treatment Effects If there is a small main treatment effect, to achieve a large beneficial effect for a subgroup, there needs to be a subgroup that receives no benefit or is even harmed by the treatment. Put simply, if there is a subgroup that does well, it must be balanced by a subgroup that does poorly.[30] This seems a very unlikely scenario when treatments are compared to a placebo, a waiting list, or no treatment, but it is more tenable when 2 active treatments are compared in comparative effectiveness trials.

Clinicians Never See Treatment Effects, Let Alone Treatment-Effect Modification It is ironic that many clinicians strongly believe in treatment-effect modification, because they never get to see a treatment effect (or its modification) in clinical practice. A treatment effect is the difference in outcome that arises from administration of treatment A compared to the outcome that arises with administration of treatment B, and is usually established from the study of large groups of patients in randomized controlled trials.[20] Potentially, one could observe treatment effects in individual patients with a single-case experimental design if one were to alternately apply treatments to a single patient, use random allocation of order of treatments, administer each treatment more than once to account for effects of time, and have wash-out periods to avoid confounding and measure outcomes closely over time. But this level of complexity is not normally what happens in clinical care, and it is so cumbersome that people generally prefer to test treatments in randomized controlled trials. The next problem would be that one would need to do this on many patients to establish differential treatment effects related to a key patient characteristic, while also accounting for effects of time—a very difficult task that is unlikely to occur in routine clinical practice.

What clinicians do see in clinical practice are treatment outcomes (ie, the outcome following delivery of a treatment), which comprise the clinical course. If they see a sufficient number of cases, they may notice patient attributes linked to different clinical courses. So, while it is plausible that experienced clinicians could discern prognostic factors, it is less likely that they could discern treatment-effect modifiers. Why, then, would clinicians strongly believe in treatment-effect modification? Two aspects of human nature that could explain this situation are that we tend to see patterns where none exist (patternicity)[27] and that we presume we have more control over events than we truly do (illusion of control).[12]

Subgroup Analysis Can Restrict Treatment Options Subgroup analyses are associated with a high risk of false-positive and false-negative results when not performed correctly.[9] They can falsely indicate that there is no treatment effect in a particular subgroup when there is a true effect,[9,30] which may restrict effective treatment options from being used by clinicians. A simulated study showed that from an overall nonsignificant result, the chance of spuriously finding at least 1 subgroup-specific test significant could be as high as 21%, and when the overall finding is significant, this can be as high as 2 in 3 tests if not tested for subgroup-treatment effect interactions.[9] That means that if there is a false-positive baseline characteristic (eg, higher score of pain) that was associated with an increased benefit from one intervention (eg, electrotherapy), then these patients could be restricted from other interventions (eg, exercise) that may be truly effective, and treatment will be restricted to this subgroup. This is also a problem with underpowered subgroup analyses.

Additionally, small groups of patients who are usually underrepresented in tri-

als are also more likely to be restricted from a potential treatment benefit. For instance, the findings from The Canadian Cooperative Study Group[11] showed that aspirin reduces the risk of continuing ischemic attacks, stroke, and death in men but not in women. This mistaken subgroup claim led to women being denied a beneficial intervention for decades, until this was demystified by subsequent studies.[4] Thus, it may not be worthwhile to risk effective interventions being restricted from clinical practice based on such spurious findings.

CONCLUSION

There is great interest in subgrouping patients with nonspecific LBP. Proponents see chances to better tailor treatments to patients based on clinical characteristics. At the same time, we must conclude that in general, the current research initiatives and achievements in this field are far from optimal and not yet ready to be implemented in clinical practice. ◉

REFERENCES

1. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J*. 2006;15 suppl 2:S192-S300. https://doi.org/10.1007/s00586-006-1072-1
2. American Physical Therapy Association. Clinical research agenda for physical therapy. *Phys Ther*. 2000;80:499-513.
3. Annadurai K, Danasekaran R, Mani G. Personalized medicine: a paradigm shift towards promising health care. *J Pharm Bioallied Sci*. 2016;8:77-78. https://doi.org/10.4103/0975-7406.171732
4. Antiplatelet Trialists' Collaboration. Collaborative overview of randomised trials of antiplatelet therapy - 1. Prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. *BMJ*. 1994;308:81-106. https://doi.org/10.1136/bmj.308.6921.81
5. BenDebba M, Torgerson WS, Long DM. A validated, practical classification procedure for many persistent low back pain patients. *Pain*. 2000;87:89-97. https://doi.org/10.1016/S0304-3959(00)00278-5
6. Bishop FL, Dima AL, Ngui J, et al. "Lovely pie in the sky plans": a qualitative study of clinicians' perspectives on guidelines for managing low back pain in primary care in England. *Spine (Phila Pa 1976)*. 2015;40:1842-1850. https://doi.org/10.1097/BRS.0000000000001215
7. Borkan JM, Koes B, Reis S, Cherkin DC. A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. *Spine (Phila Pa 1976)*. 1998;23:1992-1996.
8. Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, Erhard RE. Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. *Spine (Phila Pa 1976)*. 2006;31:623-631. https://doi.org/10.1097/01.brs.0000202807.72292.a8
9. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. *Health Technol Assess*. 2001;5:1-56.
10. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. *BMJ*. 2015;351:h5651. https://doi.org/10.1136/bmj.h5651
11. The Canadian Cooperative Study Group. A randomized trial of aspirin and sulfinpyrazone in threatened stroke. *N Engl J Med*. 1978;299:53-59. https://doi.org/10.1056/NEJM197807132990201
12. Casarett D. The science of choosing wisely—overcoming the therapeutic illusion. *N Engl J Med*. 2016;374:1203-1205. https://doi.org/10.1056/NEJMp1516803
13. Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med*. 2004;141:920-928. https://doi.org/10.7326/0003-4819-141-12-200412210-00008
14. Downie AS, Hancock MJ, Rzewuska M, Williams CM, Lin CW, Maher CG. Trajectories of acute low back pain: a latent class growth analysis. *Pain*. 2016;157:225-234. https://doi.org/10.1097/j.pain.0000000000000351
15. Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. *Spine (Phila Pa 1976)*. 2003;28:1363-1371; discussion 1372. https://doi.org/10.1097/01.BRS.0000067115.61673.FF
16. Fritz JM, Lindsay W, Matheson JW, et al. Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. *Spine (Phila Pa 1976)*. 2007;32:E793-E800. https://doi.org/10.1097/BRS.0b013e31815d001a
17. GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. 2016;388:1603-1658. https://doi.org/10.1016/S0140-6736(16)31460-X
18. Haskins R, Cook C. Enthusiasm for prescriptive clinical prediction rules (eg, back pain and more): a quick word of caution. *Br J Sports Med*. 2016;50:960-961. https://doi.org/10.1136/bjsports-2015-095688
19. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*. 2011;378:1560-1571. https://doi.org/10.1016/S0140-6736(11)60937-9
20. Kamangar F. Effect modification in epidemiology and medicine. *Arch Iran Med*. 2012;15:575-582.
21. Kamper SJ, Maher CG, Hancock MJ, Koes BW, Croft PR, Hay E. Treatment-based subgroups of low back pain: a guide to appraisal of research studies and a summary of current evidence. *Best Pract Res Clin Rheumatol*. 2010;24:181-191. https://doi.org/10.1016/j.berh.2009.11.003
22. Kent P, Keating J. Do primary-care clinicians think that nonspecific low back pain is one condition? *Spine (Phila Pa 1976)*. 2004;29:1022-1031.
23. Kent P, Keating JL. Classification in nonspecific low back pain: what methods do primary care clinicians currently use? *Spine (Phila Pa 1976)*. 2005;30:1433-1440.
24. Kent P, Keating JL, Leboeuf-Yde C. Research methods for subgrouping low back pain. *BMC Med Res Methodol*. 2010;10:62. https://doi.org/10.1186/1471-2288-10-62
25. Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J*. 2010;19:2075-2094. https://doi.org/10.1007/s00586-010-1502-y
26. Mistry D, Patel S, Hee SW, Stallard N, Underwood M. Evaluating the quality of subgroup analyses in randomized controlled trials of therapist-delivered interventions for nonspecific low back pain: a systematic review. *Spine (Phila Pa 1976)*. 2014;39:618-629. https://doi.org/10.1097/BRS.0000000000000231
27. Nickerson RS. Confirmation bias: a ubiquitous phenomenon in many guises. *Rev Gen Psychol*. 1998;2:175-220.
28. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med*. 1992;116:78-84. https://doi.org/10.7326/0003-4819-116-1-78
29. Redekop WK, Mladsi D. The faces of personalized medicine: a framework for understanding its meaning and scope. *Value Health*. 2013;16:S4-S9. https://doi.org/10.1016/j.jval.2013.06.005
30. Rothwell PM. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*. 2005;365:176-186. https://doi.org/10.1016/

S0140-6736(05)17709-5
31. Saragiotto BT, Maher CG, Moseley AM, et al. A systematic review reveals that the credibility of subgroup claims in low back pain trials was low. *J Clin Epidemiol*. 2016;79:3-9. https://doi.org/10.1016/j.jclinepi.2016.06.003
32. Slamon D, Eiermann W, Robert N, et al. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med*. 2011;365:1273-1283. https://doi.org/10.1056/NEJMoa0910383
33. Stanton TR, Hancock MJ, Maher CG, Koes BW. Critical appraisal of clinical prediction rules that aim to optimize treatment selection for musculoskeletal conditions. *Phys Ther*. 2010;90:843-854. https://doi.org/10.2522/ptj.20090233
34. Sun X, Briel M, Busse JW, et al. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. *BMJ*. 2012;344:e1553. https://doi.org/10.1136/bmj.e1553
35. Sun X, Ioannidis JP, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. *JAMA*. 2014;311:405-411. https://doi.org/10.1001/jama.2013.285063
36. van Tulder M, Becker A, Bekkering T, et al. Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J*. 2006;15 suppl 2:S169-S191. https://doi.org/10.1007/s00586-006-1071-2