General Discussion

The aim of this thesis was to elucidate the current role of paracetamol in the treatment of low back pain (LBP) in primary care. In order to do so, we first compared the recommendations on the pharmacological management of LBP in national clinical practice guidelines. Subsequently, we intended to reproduce the results of the Paracetamol for Acute Low Back Pain (PACE) trial (1) in a follow-up clinical trial (i.e. the PACE-plus trial), but had to discontinue this trial due to insufficient patient recruitment. Finally, we conducted three secondary analyses using data collected in the PACE trial: we tested the inferential reproducibility of the conclusions drawn in PACE, we investigated the efficacy of paracetamol in acute LBP among participants who complied with regular paracetamol treatment, and we explored the association between reporting adverse events of paracetamol and outcomes of LBP. In the current chapter, we place the most important findings from these research projects in the context of prior knowledge and discuss the methodological limitations of this thesis. We end this chapter by discussing the implications of these findings for clinical practice and future research.

8.1 INTERPRETATION OF THE PRINCIPAL FINDINGS IN THE CONTEXT OF PRIOR KNOWLEDGE

8.1.1 Variation in guideline recommendations for the pharmacological management of low back pain

In **Chapter 2**, we presented the results of our systematic literature review of recommendations on the pharmacological management of LBP. The most striking result of this review was the difference between the analgesics of first choice of the eight recent national clinical practice guidelines. Four out of eight guidelines (Australia, Canada, Denmark and the Netherlands (2-5)) still recommend the prescription of paracetamol for LBP, while the other four guidelines (Belgium, Germany, Britain and the US (6-9)) recommend non-steroidal anti-inflammatory drugs (NSAIDs) as first-choice analgesic; furthermore, the American guideline offers a choice between NSAIDs and skeletal muscle relaxants.

There may be a number of explanations why the results from the PACE trial (1) have not been taken up by four recent guidelines. First of all, policymakers behind the guidelines may have considered the results of a single randomized-controlled trial (RCT) as insufficient evidence to change their recommendations. Second, the lack of a safe and effective alternative could have played a role; in fact, NSAIDs, i.e. the next step on the WHO pain ladder (10), are contra-indicated in many primary care patients (for instance due to gastro-intestinal or cardiovascular comorbidity (11)). More importantly, both the review by Machado and colleagues (12) and the upcoming revision of the Cochrane systematic review regarding the efficacy of NSAIDs for acute LBP (Van der Gaag ***et al***, submitted work) question whether the difference between NSAIDs and placebo is clini-

cally relevant. For pain intensity, Machado and colleagues found a mean difference of 9.2 on a 100-point scale in favor of NSAIDs up to 2 weeks of follow-up, which is very similar to Van der Gaag and colleagues' mean difference of 7.3 up to 3 weeks of follow-up. For physical function, the two meta-analyses used different scales: Machado and colleagues found a mean difference of 8.1 points on a 100-point scale in favor of NSAIDs up to 2 weeks of follow-up, where Van der Gaag and colleagues found a mean difference of 2.02 point on the 24-point Roland-Morris Disability Questionnaire (RMDQ) up to 3 weeks of follow-up. These mean differences correspond to a 20% extra improvement in people receiving NSAIDs as compared to people receiving placebo; this generally considered to be the smallest worthwhile effect (13, 14). Furthermore, in head-to-head comparisons of NSAIDs and paracetamol, it was found that these medicines are equally effective for pain intensity and physical functioning in patients with acute LBP, while NSAIDs were associated with more adverse effects as compared to paracetamol (15-19). Third, the four mentioned guidelines may not have implemented the PACE results because of the safety profile of paracetamol, which is still perceived as relatively safe (20). However, results from observational research have shown that this may not be justified (21); paracetamol overdose is the number one cause of acute liver failure worldwide (22), and patients taking paracetamol have an increased risk of gastro-intestinal, cardiovascular and renal AEs (21). Finally, the results of PACE have been challenged as due to non-compliance to treatment (23, 24); uncertainty regarding the efficacy of paracetamol in compliers may have been a reason for policymakers not to change the existing recommendations regarding this drug. This uncertainty is addressed in detail in **Chapter 6**.

At first glance, it seems counterintuitive that only the American and Canadian guideline recommend skeletal muscle relaxants for acute LBP as these drugs could provide a clinically relevant effect on pain on the short term (9). Abdel Shaheed's meta-analysis found a mean difference of 21.3 points on a 100-point scale in favor of skeletal muscle relaxants at 3 months follow-up (25). The main reason why the other guidelines did not make any recommendations about skeletal muscle relaxants may be because they are not widely available outside North America. This begs the question of whether we are missing out on these drugs in Europe or not. Surprisingly, the correct answer to this question may be that we are not. Although the effect on pain was considered to be clinically relevant, the effect on disability was not: the mean difference was 6.5 points on a 100-point scale in favor of skeletal muscle relaxants (25). Furthermore, a recent RCT, which could not yet be included in Abdel Shaheed's meta-analysis, that compared a combination of ibuprofen and skeletal muscle relaxants (baclofen, metaxalone or tizanidine) to a combination of ibuprofen and placebo concluded that adding skeletal muscle relaxants to ibuprofen did not improve pain or physical functioning after one week of follow-up in patients that visited the emergency department for acute LBP (26). Moreover, skeletal muscle relaxants are associated with unpleasant adverse events (sedation, nausea, vomiting, vision

problems, headaches and dizziness), a potential for abuse and dependency and substantial costs (27). As there is still uncertainty about the balance between benefits and harms of skeletal muscle relaxants, more research is needed before adequate guideline recommendations can be made about the use of these drugs for the treatment of acute LBP.

The most important conclusion from the systematic review presented in **Chapter 2** is that although guidelines are universally moving away from recommending pharmacotherapy, there is currently no consensus regarding the analgesic of first choice in case patients do require medication. From the meta-analyses that we considered as best available evidence for the efficacy on drug efficacy, it can be concluded that all pharmacological treatments only have small to moderate short-term effects for non-specific LBP at best. Considering that for the majority of patients, the natural course of acute LBP is favorable (28), the best treatment of non-specific LBP in primary care may therefore not be the one with the best efficacy, but the one with the least side effects. Following this principle, the American guideline recommends nonpharmacological treatments (the efficacy of which is also small to moderate at best) rather than pharmacological treatment for recent-onset non-specific LBP (9). It is now up to policymakers from other countries whether they choose to follow this example in the upcoming revisions or not.

Limitations

The most important limitation of **Chapter 2** is that it was a narrative review, in which a number of arbitrary decisions were made in the methods. An example is the criterion to include only 'recent' clinical practice guidelines, which we defined as published after January 1st, 2016. However, as it was our aim to compare recent guidelines to the best available evidence, a limit would have been necessary in any case, and the definition of recentness would always have been subjective. Another design choice that could have influenced the results was restriction of languages to English, German or Dutch. Because of these criteria, we could have missed clinical practice guidelines. However, when we compare the guidelines included in our review to another recent overview aiming to investigate the consensus among clinical practice guidelines, we can see we did not miss any guideline because of language restrictions (29). Finally, it is difficult to judge what the 'best available evidence' is. As new studies are constantly published, we decided the best available evidence regarding the efficacy of pharmacological treatments would be in the most recent systematic reviews and meta-analyses, which constitute the top of the evidence pyramid (30). Alternatively, we could also have chosen to include only adequately powered RCTs with low risk of bias (31). Irrespective of these limitations, this review provided a valuable overview to provide context for other projects in this thesis.

8.1.2 The development and discontinuation of the PACE Plus trial

A protocol for a clinical trial to follow-up the Australian PACE trial (1) was presented in **Chapter 3**. With the PACE Plus trial, we attempted to reproduce the results of PACE by comparing paracetamol and placebo for the treatment of acute non-specific LBP in primary care; simultaneously, we also intended to compare paracetamol to diclofenac, the most prescribed NSAID in the Netherlands (32), and one of the most used NSAIDs worldwide (33). Finally, we planned to compare the combination of advice and medication (paracetamol or diclofenac) to advice alone. Together, these comparisons were designed to answer three important research questions that arose after publication of the results of the PACE trial. First, can the results of the PACE trial be reproduced? Second, considering that NSAIDs have not been demonstrated to be consistently superior to paracetamol (15-19), which in turn has been shown not to have a clinically relevant effect on LBP as compared to placebo (1), what is the comparative effectiveness of paracetamol, diclofenac and placebo for acute LBP? Third, would treating patients with acute LBP with advice and reassurance only be inferior to treating these patients with paracetamol or diclofenac?

After 6 months of recruitment in the PACE Plus trial, only four of the required 800 patients with LBP had been recruited, leading to the discontinuation of the RCT. In order to investigate the underlying reasons for termination of this RCT, we conducted a survey among local research coordinators of the participating GP practices; results of this survey have been shown in **Chapter 4**. GPs mentioned an insufficient number of patients meeting the study's eligibility criteria, lack of time in daily practice, and different patient expectations as the three main reasons for failed patient recruitment in PACE Plus (in order of descending number of comments); together, these three reasons formed over half of all reasons reported in the survey (48 out of 81 reported reasons). In a systematic literature review investigating factors that limit the progress of RCTs, common barriers to participation in clinical trials were very similar to what was found in the survey in **Chapter 4** (34).

The reduced number of patients with acute LBP seeking help in general practice when compared to the reported incidence of acute LBP in the Netherlands (35) could be explained by Lasagna's law (36, 37), the observation that once a trial starts, the number of available patients is between a tenth and a third of what was originally expected by the researchers (37). This was first described by American physician Louis Lasagna, who defined this phenomenon as "the incidence of patient availability sharply decreases when a clinical trial begins" (38). Explanations for Lasagna's law may be that researchers and clinicians overestimate the number of available patients before the study (for instance because not all patients with a new episode of disease are willing to be randomized), or that clinicians have insufficient time to recruit available cases. Another reason for this suspected drop in incidence could be due to a true decline in GP visits between the mea-

surement of the incidence of LBP (which was in 2012) and the start of the PACE Plus trial (which was in September of 2016). An explanation of such a development could be the launch of the patient information website of the Dutch College of General Practitioners, which happened in March 2012 (39, 40). In 2014, this website already had 2.9 million unique views per month, and a decline of 12% in consultations in general practice was attributed to the website launch; it is highly likely that this reduction in consultations has become even larger in the five years that have passed since then, as the number of unique views per month has increased to 4.6 million in June 2019 (39, 41). The recommendations on the web page about LBP are mainly aimed at improving self-management of complaints: they focus on the favorable prognosis of LBP, the limited benefits of medication and imaging and the necessity to remain active. This information could have directly contributed to one of the other reasons for insufficient patient recruitment mentioned in the survey, namely the effective self-management of LBP. Another explanation of a true decline in GP visits of patients with LBP could be the increasing popularity of direct access to physiotherapy; in 2017, 56% of patients used direct access to physiotherapy as compared to 35% in 2009 (42, 43). Apart from the change in number of patients visiting their GP with recent onset LBP, there may also have been a change in type of patients with LBP presenting in general practice. Patients who recovered using self-management skills only might not visit their GP, whereas the patients seeking care with their GP may have had LBP for a longer period, with greater limitations of their usual daily activities and with unsatisfactory results using over-the-counter medication such as paracetamol and NSAIDs. If this is the case, it seems logical that these patients have different expectations than participating in a trial that offers them exactly the same interventions they have been using for a number of weeks. In light of this suspected changing population of patients presenting to general practice with recent onset LBP, recruiting incident cases of LBP from Dutch general practice for research into first-choice interventions may remain a challenge. However, conducting research in clinical practice is not just a challenge in the Netherlands, but also in other countries such as the United States and the United Kingdom (44); the feasibility of such a study is thus not automatically guaranteed if it were to be conducted in another country.

The second most frequent reason (i.e. lack of time in general practice) seems to reflect the current state of general practice in the Netherlands. Because of changes in the national health care system, the range of tasks of the Dutch GP has vastly increased, as has the related administrative workload; similar increases in workload have been reported in England, which has a similar organization structure of primary care (45). In 2015, this even led to an action group presenting a manifesto to the House of Representatives of the Netherlands signed by two-thirds of all Dutch GPs (a total of 7800 signatures), in which they asked for health care system reforms (46). In an international comparison of the workload of general practitioners, it was found that Dutch GPs spent a large percent-

age of time on tasks that were not directly patient-related, such as administration, when compared to GPs from other countries (47). Although some positive changes have been made in recent years, such as the decrease in number of patients per practice, conducting clinical trials in Dutch general practice (for any clinical condition) will probably remain difficult in the foreseeable future.

GPs could be encouraged to participate in research by allowing for adequate reimbursement of their invested time. Even though a Dutch study investigating the factors related to success and failure of patient recruitment did not identify GP reimbursement as a factor influencing trial success (36), the comparison between PACE Plus and the original PACE trial revealed this may have been an important difference between the Australian and the Dutch trial. However, this would require an increase in the budget of research projects, which would probably only be possible in collaborations with the pharmaceutical industry. Contrary to the reasoning in the Cochrane Risk of Bias assessment tool (31), industry involvement does not always have to be a cause for concern, as long as corporate sponsors are not involved in study design and analysis and interpretation of results, as demonstrated by the PACE trial (1); however, it will very likely be difficult to obtain industry funding under these terms in practice. As an alternative solution, the Australian system for mandatory continuing education, in which part of the points obligatory for re-registration as a GP can only be earned by participating in research, could be implemented in the Netherlands.

Limitations

The most significant limitation of the PACE Plus trial was by far its feasibility. With different study design choices, the trial may have been more likely to succeed. A number of general recommendations for future research have already been made in **Chapter 3**. As the specific research questions of PACE Plus remain relevant but unanswered, an alternative approach to PACE Plus is presented in **Section 8.3**.

From a technical perspective, a limitation of the PACE Plus trial was the fact that it is not strictly speaking a results reproduction study according to Goodman's new lexicon of research reproducibility (48). Goodman describes results reproduction as the collection of new data in the same population and consequently analyzing this data using the same analysis plan (48). Although the first criterion was met (both trials were set in primary care, albeit on opposite sides of the globe), the second was not. While the original PACE trial had three treatment groups (paracetamol taken regularly, paracetamol as-needed for pain and placebo), PACE Plus had four (paracetamol, diclofenac, placebo and advice only), meaning a different analysis plan was needed. Furthermore, not all outcomes were identical between the two trials, the most notable difference being the measurement of health-related quality of life (HRQoL), which we intended to record using the EuroQol Group 5 Dimensions, 5 Level Questionnaire (EQ-5D-5L (49)) rather than the Short Form

12 (SF-12 (50)) that was used in the original trial. These instruments provide different HRQoL scores (i.e. the EQ-5D-5L provides overall utility and visual analogue scale scores, whereas the SF-12 provides a physical and a mental summary score) that make a comparison very difficult to do (Chiarotto 2018 Pain). But although PACE Plus was not strictly speaking a results reproduction study, data on pain intensity, disability and time until recovery from LBP from PACE Plus could have been combined with data of PACE in an individual patient data (IPD) meta-analysis (51), had the PACE Plus trial been completed.

An important technical consideration about the survey in **Chapter 3** was the way the survey was conducted. Although the response rate among local research coordinators was high (92%), it is debatable whether these 33 research coordinators were representative of all 96 GPs that participated in the trial. Furthermore, the survey was not a structured combination of open- and close-ended questions, as it solely consisted of a single open question. By performing a more elaborate survey among all participating GPs rather than only the local coordinators, a more valid picture could have been obtained; however, given the high workload of GPs, a more elaborate survey would likely have had a much lower response rate and thus a poorer representation of the participating clinicians.

8.1.3 Secondary analyses of data collected in the PACE trial

Three projects presented in this thesis were based on secondary statistical analyses of data collected in the PACE trial (1). Our first re-analysis of PACE focused on the reproducibility of the knowledge claims made in the original data analysis. One of the reasons for failed patient recruitment in PACE Plus mentioned by four GPs in the survey in **Chapter 4** was that the research question of the PACE Plus trial was irrelevant for clinical practice, since the original PACE trial had already sufficiently investigated efficacy of paracetamol for LBP. However, reproducibility is one of the cornerstones of scientific research (52) and many scientific claims are found not to be reproducible (53-55). In **Chapter 5**, we presented the results of the first independent inferential reproducibility study in the LBP research field. This study focused on the reproduction of the causal inferences of the PACE trial for the core outcome domains of LBP: pain intensity, physical functioning and HRQoL (56). We analyzed the data in the PACE trial with an independent team using the pre-defined and published statistical analysis plan from the PACE Plus trial; the original PACE trial authors had no influence on the aim, methods and conclusions of this study and effectively gave us “carte blanche” to conduct this analysis. In the reproducibility study, paracetamol had no effect on the core outcomes when compared to placebo.

In our second analysis, we investigated the efficacy of paracetamol among participants who complied to treatment. As stated earlier, although it was demonstrated that paracetamol had no overall effect on outcomes of acute LBP when compared to placebo (1), it was unclear if there was a difference between paracetamol and placebo in compli-

ers to the treatment regimen; this may have played a role in the observation that four recent guidelines did not change their recommendation regarding paracetamol for LBP. In **Chapter 6**, we showed that paracetamol was not more effective than placebo for acute LBP, regardless of the definition of compliance or follow-up period, using a complier average causal effects (CACE) analysis.

In **Chapter 7**, we presented our secondary analysis of PACE, in which we looked into the association of reporting adverse events (AEs) and on the one hand, baseline characteristics and on the other hand, outcomes of LBP. Baseline characteristics that were associated with reporting AEs were older age, more days since the onset of pain, increased feelings of depression. The strongest association was found for the use of medicines for a health problem other than LBP (odds ratio 1.42, 95% confidence interval 1.07 – 1.88), suggesting that not all reported AEs were related to taking trial medication. No association was found between reporting AEs and the core outcome domains of LBP at follow-up (56).

These findings should be interpreted in the context of the original (primary) analysis of the PACE trial (1). The fact that the results from the independent inferential reproduction analysis are consistent with the original results of PACE, even though a different approach to the statistical analysis was used, strengthens the conclusions regarding the lack of efficacy of paracetamol for acute LBP. Our second analysis represented only the second time the CACE analysis technique was used in the LBP research field (57). The findings of this study extend the message of the original analysis of PACE and form a strong appeal to clinicians and policymakers to reconsider their endorsement of paracetamol for the treatment of acute LBP. Finally, our findings in the AEs analysis suggested that if LBP-patients decide to take paracetamol anyway and consequently experience AEs, overall this is not associated with less favorable outcomes of LBP. If we combine all these results, the bottom line seems to be that taking paracetamol has very little influence (neither negative nor positive) on the outcomes of acute LBP.

Limitations

The most important limitation of the studies presented in **Chapters 5, 6 and 7** was that existing data regarding the efficacy of paracetamol for acute LBP was used; no new data was collected during these studies. An additional disadvantage of this is that the available data was not intended to conduct a CACE analysis on or to investigate the association between reporting AEs of paracetamol and outcomes of LBP. The accuracy of the CACE analysis could have been improved by including a measurement for the likelihood of compliance in the baseline questionnaire (58). An important limitation of the AEs analysis is there was no verification whether or not reported AEs were related to taking study medicines in PACE.

8.2 IMPLICATIONS FOR CLINICAL PRACTICE

In spite of the fact that the evidence for a lack of efficacy of paracetamol for acute LBP has been strengthened by the studies in this thesis, the most important piece of the puzzle is still missing: results reproduction of the PACE trial. However, irrespective of the efficacy of paracetamol, it seems that GPs may have taken a wrong turn somewhere in the past decades when it comes to the management of LBP, considering the evidence presented in **Chapter 2**. We still tend to focus on pain intensity and the treatment of pain until a pain level of zero is reached, rather than focusing on the influence that LBP has on the daily activities of our patients and the treatment until an acceptable level of functioning is achieved (56, 59-61).

Although the treatment of pain in evidence-based medicine is currently strongly associated with the prescription of medication, we must not forget that there are other therapeutic options in clinicians' toolkits. Treatment options for acute LBP that are feasible, affordable and available today include superficial heat, massage, or exercise, irrespective of their specific efficacy; research investigating placebo-interventions suggests that for patient reported outcomes, almost any intervention is better than no intervention at all (62). Furthermore, even open-label placebo interventions have demonstrated clinically relevant effects on pain, given that they are provided in a positive context (63). Ideally, these suggested non-pharmacological treatments should therefore be wrapped in the best possible 'therapeutic envelope', which is also known as the patient-provider interaction. In a time when it has been shown that most pharmacological treatments have no clinically relevant effects, larger benefits may be expected from maximizing contextual effects than from the development of new drugs.

Instead of a clear and unambiguous recommendation for clinical practice, this thesis provides an ethical dilemma: should clinicians still prescribe paracetamol to patients with acute LBP, now that we know that even in patients who comply with the treatment regimen, it has no effect when compared to placebo? On the one hand, one could argue they should; placebos are associated with effects on patient reported outcomes (64) and one could argue that although paracetamol is definitely not harmless (21), of all the analgesics, it arguably has the most favorable safety profile. If clinicians stop prescribing paracetamol, many patients that require pain medication will be prescribed NSAIDs or even opioids instead. In such a scenario, are we not better off if we just continue to prescribe paracetamol? On the other hand, in conventional medicine, it is not acceptable to prescribe placebo tablets (which have no characteristic effect, but no adverse effects either). If paracetamol is prescribed solely for the purpose of prescribing a placebo, then are we not breaking our own rules? Furthermore, if we know that the effects of an intervention is based on the placebo effect, but we continue to use the intervention anyway, then what is the difference between conventional evidence-based medicine and

what we call alternative medicine (65)? And if we purposely prescribe paracetamol as a placebo, how sure are we that this placebo effect is clinically worthwhile to LBP patients (13, 14)?

8.3 IMPLICATIONS FOR FUTURE RESEARCH

The discontinuation of the PACE Plus trial left a legacy of unanswered research questions. First, the results of the PACE trial should be reproduced. Second, the relative efficacy of paracetamol as compared to that of NSAIDs and other medicaments and non-pharmacological interventions remains unclear. Third, we still have limited and contradictory evidence regarding the magnitude and clinical relevance of the placebo effect in LBP when compared to no treatment (or waiting list controls).

Reproduction of the results of PACE will be challenging in practice, as it requires the collection of new data. Currently, the recruitment of incident cases of acute LBP in Dutch general practice does not seem realistic, given the amount of pressure GPs are already under due to their normal workload. However, we know that the one-year risk of recurrence of acute LBP is 33% (66, 67). An RCT nested in a cohort of prevalent cases of LBP may therefore be much more feasible. Instead of burdening GPs with patient recruitment, cohort participants could be recruited through large population databases, such has the Integrated Primary Care Information (IPCI) database (68), a large database with 1.5 million primary care patient records from the South of the Netherlands. For instance, patients for whom an ICPC-code L03 (non-radiating LBP) was registered in the past year could be approached via post to participate in the cohort, and could subsequently be instructed to contact the research department as soon as they experience a new episode of LBP. The most feasible approach would be to then randomize these participants to regular paracetamol or placebo, thus answering our research question. Another option would be to use a large observational dataset to answer this research question, given that confounding can be adequately assessed and corrected for (which is the main disadvantage of non-randomized studies). In this scenario, the PACE Plus design could be used as a target trial protocol (69).

Although our experience with PACE Plus suggests the likelihood of failure may increase by attempting to answer multiple research questions in one study (in RCTs, do not try to kill two birds with one stone), it would theoretically be possible in a cohort-nested RCT design to simultaneously investigate the second unanswered research question of PACE Plus as well. In this scenario, patients could be randomized to receive non-pharmacological treatment only (which could consist of advice and reassurance only, superficial heat and or/exercises), or to receive a combination of non-pharmacological treatment and medication. In the latter group, patients could be allocated to different treatment options

(such as paracetamol, ibuprofen or placebo) using a second randomization procedure. A downside of this approach is that it is very costly and labor-intensive. An alternative approach to comparing paracetamol to other pharmacological and non-pharmacological interventions would be by conducting a network meta-analysis (70, 71), in which direct and indirect evidence is combined to reconstruct a comparison between interventions that have never been tested head-to-head in an RCT. Protocols have already been developed for the conduct of such network meta-analyses in order to compare the effects and safety of paracetamol, NSAIDs and opioids for chronic LBP (72) and to compare the effects of all noninvasive treatments of LBP (73).

Apart from ambiguous evidence regarding the comparative efficacy of paracetamol and NSAIDs, it is unclear how acceptable these treatments currently are to patients with LBP. NSAIDs in particular have been in the media in a negative context during recent years (74-76). Observational studies have been conducted for physiotherapy and NSAIDs to investigate when the effects of these interventions become worthwhile to LBP patients (13, 14). Given the fact that over the last decade, many new studies been published about both the efficacy and the safety of paracetamol and NSAIDs (12, 21, 77), it is important to update the current knowledge on when these medicines are clinically worthwhile to patients with LBP in order to place their efficacy into context.

The PACE Plus trial would have been the first RCT in the LBP research field to directly compare a combination of advice and blinded placebo tablets to advice only. The only similar published result would be a comparison between open-label placebo and treatment as usual in people with chronic LBP (63). In this study, open label placebo pills were found to have a significant and possibly clinically relevant effect on both pain and disability after 3 weeks of follow-up when compared to treatment as usual (63). However, as this study used open-label placebo pills which were provided in a positive context, it is unclear whether the magnitude of the effects found in this study reflects the placebo effect (in this case: the effect of a placebo intervention (78)) associated with medication given to people with LBP. The reason why knowledge about the effect size associated with a placebo intervention is so important, is because active interventions are compared to these placebo-interventions in clinical trials in order to determine the efficacy of these interventions. If the effect of placebo interventions compared to no treatment (or for instance a waiting list) is not clinically relevant, then this has implications for the interpretation of the results of placebo-controlled trials as well. Currently, the existence and clinical relevance of the effects of placebo interventions is a topic of debate in the LBP research field: on the one hand, it was found in a meta-analysis that the pooled magnitude of placebo effects for pain is very small (3.2 points on a 100-point pain scale) (79) and on the other hand, it was concluded from a systematic review that placebo tablets could have a clinically meaningful effect on pain in people with non-specific LBP (80). However, the search strategies for these studies are now over ten years old, so many new

RCTs comparing placebo interventions to no treatment may have been conducted in the meantime (81). As stated in **Section 8.2**, another important topic in placebo research is the optimization of the patient-provider interaction (82). New research in this area should mainly focus on the development and implementation of specific evidence-based strategies for general practice. By using the doctor as a medicine, we may be able to avoid the pills (83).

REFERENCES

1. Williams CM, Maher CG, Latimer J, McLachlan AJ, Hancock MJ, Day RO, et al. Efficacy of paracetamol for acute low-back pain: a double-blind, randomised controlled trial. Lancet. 2014;384(9954):1586-96.
2. NSW Agency for Clinical Innovation. Management of people with acute low back pain: model of care. Chatswood; NSW Health. 2016:39 p.
3. Institute of Health Economics Alberta Canada. Evidence-Informed Primary Care Management of Low Back Pain. wwwiheca or wwwtopalbertadoctorsorg. 2017.
4. Stochkendahl MJ, Kjaer P, Hartvigsen J, Kongsted A, Aaboe J, Andersen M, et al. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. Eur Spine J. 2018;27(1):60-75.
5. Bons S.C.S., Borg M.A.J.P., Van den Donk M., Koes B.W., Kuijpers T., Ostelo R.W.J.G., et al. The revised Dutch College of General Practitioners (NHG) practice guideline on 'Non-specific Low Back Pain'(in Dutch). Huisarts Wet. 2017;60(2).
6. Van Wambeke P, Desomer A, Ailliet L, ABerquin A, Demoulin C, Depreitere B, et al. Low Back Pain And Radicular Pain: Assessment And Management. Belgian Health Care Knowledge Centre. 2017.
7. Chenot JF, Greitemann B, Kladny B, Petzke F, Pfingsten M, Schorr SG. Non-Specific Low Back Pain. Dtsch Arztebl Int. 2017;114(51-52):883-90.
8. Bernstein IA, Malik Q, Carville S, Ward S. Low back pain and sciatica: summary of NICE guidance. BMJ. 2017;356:i6748.
9. Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of P. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. Annals of Internal Medicine. 2017;166(7):514-30.
10. World Health Organization. Cancer Pain Relief. 1986:Geneva: WHO.
11. Schmidt M, Sorensen HT, Pedersen L. Diclofenac use and cardiovascular risks: series of nationwide cohort studies. BMJ. 2018;362:k3426.
12. Machado GC, Maher CG, Ferreira PH, Day RO, Pinheiro MB, Ferreira ML. Non-steroidal anti-inflammatory drugs for spinal pain: a systematic review and meta-analysis. Ann Rheum Dis. 2017;76(7):1269-78.
13. Ferreira ML, Ferreira PH, Herbert RD, Latimer J. People with low back pain typically need to feel 'much better' to consider intervention worthwhile: an observational study. Aust J Physiother. 2009;55(2):123-7.
14. Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Grotle M, et al. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit-harm trade-off study. Journal of Clinical Epidemiology. 2013;66(12):1397-404.
15. Roelofs P, Deyo R, Koes B, Scholten R, van Tulder M. Non-steroidal anti-inflammatory drugs for low back pain. Cochrane Database Syst Rev. 2008(1):CD000396.
16. Evans DP, Burke MS, Newcombe RG. Medicines of choice in low back pain. Curr Med Res Opin. 1980;6(8):540-7.
17. Milgrom C, Finestone A, Lev B, Wiener M, Floman Y. Overexertional lumbar and thoracic back pain among recruits: a prospective study of risk factors and treatment regimens. J Spinal Disord. 1993;6(3):187-93.

18. Nadler SF, Steiner DJ, Erasala GN, Hengehold DA, Hinkle RT, Beth Goodale M, et al. Continuous low-level heat wrap therapy provides more efficacy than Ibuprofen and acetaminophen for acute low back pain. Spine (Phila Pa 1976). 2002;27(10):1012-7.

19. Wiesel SW, Cuckler JM, Deluca F, Jones F, Zeide MS, Rothman RH. Acute low-back pain. An objective analysis of conservative therapy. Spine (Phila Pa 1976). 1980;5(4):324-30.

20. Herndon CM, Dankenbring DM. Patient perception and knowledge of acetaminophen in a large family medicine service. J Pain Palliat Care Pharmacother. 2014;28(2):109-16.

21. Roberts E, Delgado Nunes V, Buckner S, Latchem S, Constanti M, Miller P, et al. Paracetamol: not as safe as we thought? A systematic literature review of observational studies. Ann Rheum Dis. 2016;75(3):552-9.

22. Bunchorntavakul C, Reddy KR. Acetaminophen (APAP or N-Acetyl-p-Aminophenol) and Acute Liver Failure. Clin Liver Dis. 2018;22(2):325-46.

23. Koes BW, Enthoven WT. Do patients with acute low-back pain need paracetamol? Lancet. 2014;384(9954):1556-7.

24. Kress HG, Untersteiner G. Clinical update on benefit versus risks of oral paracetamol alone or with codeine: still a good option? Curr Med Res Opin. 2017;33(2):289-304.

25. Abdel Shaheed C, Maher CG, Williams KA, McLachlan AJ. Efficacy and tolerability of muscle relaxants for low back pain: Systematic review and meta-analysis. Eur J Pain. 2017;21(2):228-37.

26. Friedman BW, Irizarry E, Solorzano C, Zias E, Pearlman S, Wollowitz A, et al. A Randomized, Placebo-Controlled Trial of Ibuprofen Plus Metaxalone, Tizanidine, or Baclofen for Acute Low Back Pain. Ann Emerg Med. 2019.

27. Witenko C, Moorman-Li R, Motycka C, Duane K, Hincapie-Castillo J, Leonard P, et al. Considerations for the appropriate use of skeletal muscle relaxants for the management of acute low back pain. P T. 2014;39(6):427-35.

28. Vasseljen O, Woodhouse A, Bjorngaard JH, Leivseth L. Natural course of acute neck and low back pain in the general population: the HUNT study. Pain. 2013;154(8):1237-44.

29. Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CC, Chenot JF, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. Eur Spine J. 2018;27(11):2791-803.

30. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21(4):125-7.

31. Higgins JPT, Sterne JAC, Savović J, Page MJ, Hróbjartsson A, Boutron I, et al. A revised tool for assessing risk of bias in randomized trials. Chandler J, McKenzie J, Boutron I, Welch V (editors) Cochrane Methods Cochrane Database of Systematic Reviews. 2016(10).

32. Foundation for Pharmaceutical Statistics. Facts and Figures 2016 (in Dutch). The Hague, Netherlands. 2016.

33. Conaghan PG. A turbulent decade for NSAIDs: update on current concepts of classification, epidemiology, comparative efficacy, and toxicity. Rheumatol Int. 2012;32(6):1491-502.

34. Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, et al. Factors that limit the quality, number and progress of randomised controlled trials. Health Technol Assess. 1999;3(20):1-143.

35. Ursum J HK, Spronk I, Nielen MMJ, Davids R, Verheij RA. Hoe vaak komt rugpijn voor en bij wie? Uit: Nivel Zorgregistraties eerste lijn [internet]. 2019 [Available from: www.nivel.nl/nl/nivel-zorgregistraties-eerste-lijn/hoe-vaak-komt-rugpijn-voor-en-bij-wie.

36. van der Wouden JC, Blankenstein AH, Huibers MJ, van der Windt DA, Stalman WA, Verhagen AP. Survey among 78 studies showed that Lasagna's law holds in Dutch primary care research. Journal of Clinical Epidemiology. 2007;60(8):819-24.
37. Knottnerus JA, Tugwell P. Prevention of premature trial discontinuation: how to counter Lasagna's law. Journal of Clinical Epidemiology. 2016;80:1-2.
38. Harris E.L., Fitzgerald J.D. The principles and practice of clinical trials. Livingstone, Edinburgh/London. 1970.
39. Spoelman WA, Bonten TN, de Waal MW, Drenthen T, Smeele IJ, Nielen MM, et al. Effect of an evidence-based website on healthcare usage: an interrupted time-series study. BMJ Open. 2016;6(11):e013166.
40. Dutch College of General Practitioners. Thuisarts.nl 2019 [Available from: www.thuisarts.nl.
41. Nederlands Huisartsen Genootschap (NHG). Thuisarts.nl populair bij burgers en artsen: Nederlands Huisartsen Genootschap (NHG); 2019 [updated July 15, 2019. Available from: https://www.nhg.org/actueel/nieuws/thuisartsnl-populair-bij-burgers-en-artsen.
42. Barten DJ, Verberne L, Koppes L. Zorg door de fysiotherapeut: jaarcijfers 2013 en trendcijfers 2009-2013.: NIVEL; 2015 [updated 29 September 2014. Available from: https://www.nivel.nl/nl/publicatie/zorg-door-de-fysiotherapeut-jaarcijfers-2013-en-trendcijfers-2009-2013.
43. Van den Dool J, Schermer T. Zorg door de fysiotherapeut: jaarcijers 2017 en trendcijfers 2013-2017.: NIVEL; 2018 [updated 13 November 2018. Available from: https://www.nivel.nl/nl/publicatie/zorg-door-de-fysiotherapeut-jaarcijers-2017-en-trendcijfers-2013-2017.
44. Mapstone J, Elbourne D, Roberts I. Strategies to improve recruitment to research studies. Cochrane Database Syst Rev. 2007(2):MR000013.
45. Thompson M, Walter F. Increases in general practice workload in England. Lancet. 2016;387(10035):2270-2.
46. Visser J. Huisartsen luiden noodklok in Tweede Kamer 2015 [Available from: www.medischcontact.nl.
47. Schäfer W, Van den Berg M, Groenewegen P. Workload of Dutch general practitioners from an international perspective. Huisarts Wet. 2016;59(3):94-101.
48. Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? Sci Transl Med. 2016;8(341):341ps12.
49. Lamers LM, Stalmeier PF, McDonnell J, Krabbe PF, van Busschbach JJ. Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff (in Dutch). Ned Tijdschr Geneesk. 2005;149(28):1574-8.
50. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care. 2004;42(9):851-9.
51. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Eval Health Prof. 2002;25(1):76-97.
52. McNutt M. Journals unite for reproducibility. Science. 2014;346(6210):679.
53. Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical research. Circ Res. 2015;116(1):116-26.
54. Ioannidis JP. Acknowledging and Overcoming Nonreproducibility in Basic and Preclinical Research. JAMA. 2017;317(10):1019-20.
55. Open Science C. PSYCHOLOGY. Estimating the reproducibility of psychological science. Science. 2015;349(6251):aac4716.

56. Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, et al. Core outcome domains for clinical trials in non-specific low back pain. Eur Spine J. 2015;24(6):1127-42.

57. Knox CR, Lall R, Hansen Z, Lamb SE. Treatment compliance and effectiveness of a cognitive behavioural intervention for low back pain: a complier average causal effect approach to the BeST data set. BMC Musculoskelet Disord. 2014;15:17.

58. Stuart EA, Jo B. Assessing the sensitivity of methods for estimating principal causal effects. Stat Methods Med Res. 2015;24(6):657-74.

59. Ballantyne JC, Sullivan MD. Intensity of Chronic Pain--The Wrong Metric? N Engl J Med. 2015;373(22):2098-9.

60. Hush JM, Refshauge K, Sullivan G, De Souza L, Maher CG, McAuley JH. Recovery: what does this mean to patients with low back pain? Arthritis Rheum. 2009;61(1):124-31.

61. Darnall B. To treat pain, study people in all their complexity. Nature. 2018;557(7703):7.

62. Wechsler ME, Kelley JM, Boyd IO, Dutile S, Marigowda G, Kirsch I, et al. Active albuterol or placebo, sham acupuncture, or no intervention in asthma. N Engl J Med. 2011;365(2):119-26.

63. Carvalho C, Caetano JM, Cunha L, Rebouta P, Kaptchuk TJ, Kirsch I. Open-label placebo treatment in chronic low back pain: a randomized controlled trial. Pain. 2016;157(12):2766-72.

64. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. Cochrane Database Syst Rev. 2010(1):CD003974.

65. Harris I. Surgery, the ultimate placebo - A surgeon cuts through the evidence.: NewSouth Publishing; 2016.

66. da Silva T, Mills K, Brown BT, Herbert RD, Maher CG, Hancock MJ. Risk of Recurrence of Low Back Pain: A Systematic Review. J Orthop Sports Phys Ther. 2017;47(5):305-13.

67. Machado GC, Maher CG, Ferreira PH, Latimer J, Koes BW, Steffens D, et al. Can Recurrence After an Acute Episode of Low Back Pain Be Predicted? Phys Ther. 2017;97(9):889-95.

68. van der Lei J, Duisterhout JS, Westerhof HP, van der Does E, Cromme PV, Boon WM, et al. The introduction of computer-based patient records in The Netherlands. Annals of Internal Medicine. 1993;119(10):1036-41.

69. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.

70. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ. 2005;331(7521):897-900.

71. Dias S, Caldwell DM. Network meta-analysis explained. Arch Dis Child Fetal Neonatal Ed. 2019;104(1):F8-F12.

72. Bagg MK, McLachlan AJ, Maher CG, Kamper SJ, Williams CM, Henschke N, et al. Paracetamol, NSAIDS and opioid analgesics for chronic low back pain: a network meta-analysis. Cochrane Database of Systematic Reviews. 2018(6).

73. Gianola S, Greta C, Andreano A, DCorbetta D, Frigerio P, VPecoraro V, et al. The effectiveness of treatments for acute and sub-acute mechanical non-specific low back pain: a protocol of systematic review and network meta-analysis: PROSPERO; 2018 [Available from: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=102527.

74. Berkhout K. Die gevaarlijke pijnstiller uitbannen, dat lukt maar niet.: NRC; 2019 [Available from: https://www.nrc.nl/nieuws/2019/03/03/die-gevaarlijke-pijnstiller-uitbannen-dat-lukt-maar-niet-a3907876.

75. Van Vugt V. Gebruik diclofenac heeft meer risico dan andere NSAID's.: Huisarts en Wetenschap; 2018 [Available from: https://www.henw.org/artikelen/gebruik-diclofenac-heeft-meer-risico-dan-andere-nsaids.

76. Curfman G. FDA strengthens warning that NSAIDs increase heart attack and stroke risk.: Harvard Health Publishing; 2019 [Available from: https://www.health.harvard.edu/blog/fda-strengthens-warning-that-nsaids-increase-heart-attack-and-stroke-risk-201507138138.

77. Saragiotto BT, Machado GC, Ferreira ML, Pinheiro MB, Abdel Shaheed C, Maher CG. Paracetamol for low back pain. Cochrane Database Syst Rev. 2016(6):CD012230.

78. Hrobjartsson A. What are the main methodological problems in the estimation of placebo effects? Journal of Clinical Epidemiology. 2002;55(5):430-5.

79. Kamper SJ, Machado LA, Herbert RD, Maher CG, McAuley JH. Trial methodology and patient characteristics did not influence the size of placebo effects on pain. Journal of Clinical Epidemiology. 2008;61(3):256-60.

80. Puhl AA, Reinhart CJ, Rok ER, Injeyan HS. An examination of the observed placebo effect associated with the treatment of low back pain- a systematic review. Pain Res Manag. 2011;16(1):45-52.

81. Schreijenberg M, Luijsterburg PA, Van Trier YD, Rizopoulos D, Koopmanschap MA, Voogt L, et al. Efficacy of paracetamol, diclofenac and advice for acute low back pain in general practice: design of a randomized controlled trial (PACE Plus). BMC Musculoskelet Disord. 2017;18(1):56.

82. Chou C. Time to Start Using Evidence-Based Approaches to Patient Engagement 2018 [Available from: https://catalyst.nejm.org/evidence-based-patient-provider-communication/.

83. Van Osselen E, Helsloot R, Van der Werf G, Van Zalinge E. De dokter als medicijn: zeventig jaar timmeren aan de arts-patiëntrelatie. Huisarts Wet. 2016;59(4):176-9.